

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ»

Навчально-науковий комплекс "Інститут прикладного системного аналізу"
(повна назва інституту/факультету)

Кафедра Системного проектування
(повна назва кафедри)

«На правах рукопису»
УДК 378.1-051-31114

«До захисту допущено»

Завідувач кафедри

_____ (підпис)

_____ (ініціали, прізвище)

“ ” _____ 20__ р.

Магістерська дисертація

зі спеціальності 8.05010102 Інформаційні технології проектування
(код і назва спеціальності)

на тему: Кластеризація переміщень абітурієнтів при поданні документів до вищих навчальних закладів

Виконав: студент 6 курсу, групи ДА-31м
(шифр групи)

Патока Єгор Романович

(прізвище, ім'я, по батькові)

_____ (підпис)

Науковий керівник к.т.н., доц. Кисельов Г.Д.

(посада, науковий ступінь, вчене звання, прізвище та ініціали)

_____ (підпис)

Консультант Охорона праці

(назва розділу)

к.б.н., доц. Гусєв А.М.

(науковий ступінь, вчене звання, прізвище, ініціали)

_____ (підпис)

Рецензент Доцент кафедри Нарисної геометрії та інженерної графіки

к.т.н. Яблонський П.М.

(посада, науковий ступінь, вчене звання, прізвище та ініціали)

_____ (підпис)

Засвідчую, що у цій магістерській дисертації
немає запозичень з праць інших авторів без
відповідних посилань.

Студент _____
(підпис)

Київ – 2015 року

**Національний технічний університет України
«Київський політехнічний інститут»**

Факультет (інститут) ННК «Інститут прикладного системного аналізу»
(повна назва)

Кафедра Системного проектування
(повна назва)

Освітньо-кваліфікаційний рівень «Магістр»
(назва ОКР)

Напрямок підготовки 6.050101 Комп'ютерні науки
(код і назва)

Спеціальність 8.05010102 Інформаційні технології проектування
(код і назва)

ЗАТВЕРДЖУЮ

Завідувач кафедри

_____ А.І.Петренко
(підпис) (ініціали, прізвище)

«__» _____ 2015 р.

ЗАВДАННЯ

на магістерську дисертацію студенту

_____ Патоці Єгору Романовичу
(прізвище, ім'я, по батькові)

1. Тема дисертації: Кластеризація переміщень абітурієнтів при поданні документів до вищих навчальних закладів

науковий керівник дисертації Кисельов Геннадій Дмитрович, к.т.н, доц.
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затверджені наказом по університету від «13» лютого 2015 р. № 19/1-ст

2. Строк подання студентом дисертації 08.06.2015

3. Об'єкт дослідження: Щорічний прийом вступників до навчання у вищих навчальних закладах

4. Предмет дослідження: Кореляція між вищими навчальними закладами

5. Перелік завдань, які потрібно розробити:

1. Дослідити умови та правила прийому до ВНЗ та визначити фактори, що впливають на вибір вступником ВНЗ.
2. Сформулювати порядок проведення експериментів, що дозволять визначити ВНЗ-конкуренти та дослідити відомі методи кластерного аналізу.

3. Сформувати вимоги до системи, що буде формувати та завантажувати дані про вступників із обраного дисертантом джерела інформації.
4. Розробити систему формування та завантаження даних у БД.
5. Виконати наведені експерименти та за отриманими даними назвати ВНЗ-конкуренти для НТУУ "КПІ", розробити рекомендації щодо заходів по залученню вступників до навчання в НТУУ "КПІ".

6. Орієнтовний перелік ілюстративного матеріалу: презентація по темі «Кластеризація переміщень абітурієнтів при поданні документів до вищих навчальних закладів»

7. Консультанти розділів дисертації

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Охорона праці	доц., к.б.н. Гусев А. М		
Основна частина	доц., к.т.н. Кисельов Г.Д.		

8. Дата видачі завдання 30.09.2014

Календарний план

№ з/п	Назва етапів виконання магістерської дисертації	Строк виконання етапів магістерської дисертації	Примітка
1	Отримання завдання	30.09.2014	
2	Збір інформації	15.01.2015	
3	Аналіз вимог завдання, вибір методів і засобів розв'язання поставленої задачі	28.02.2015	
4	Дослідження умов та правил прийому до ВНЗ та визначення факторів, що впливають на вибір вступником ВНЗ	10.03.2015	
5	Формування порядку проведення експериментів, що дозволять визначити ВНЗ-конкуренти та дослідження відомих методів кластерного аналізу	20.03.2015	
6	Формування вимог до системи, що буде формувати та завантажувати дані про вступників із обраного джерела інформації	05.04.2015	
7	Розробка системи формування та завантаження даних у БД	25.04.2015	
8	Виконання наведених експериментів. Визначення конкурентів. Надання рекомендацій		
9	Розробка розділу з охорони праці	10.05.2015	
10	Оформлення дипломної роботи	31.05.2015	
11	Отримання допуску до захисту та подача роботи в ДЕК	10.06.2015	

Студент

_____ (підпис)

Є.Р. Патока

(ініціали, прізвище)

Керівник роботи

_____ (підпис)

Г.Д. Кисельов

(ініціали, прізвище)

РЕФЕРАТ

на магістерську дисертацію

виконану на тему: Кластеризація переміщень абітурієнтів при поданні документів до вищих навчальних закладів

студентом: Патокою Єгором Романовичем

Робота виконана на 129 сторінках, містить 32 ілюстрації, 14 таблиць. При підготовці використовувалася література з 18 різних джерел.

Актуальність

Однією з невід'ємних частин життєвого циклу будь-якого ВНЗ України є щорічна вступна кампанія, метою якої є набір студентів на денну, вечірню або заочну форму навчання відповідно до обсягу державного замовлення, а також на навчання на контрактній основі.

Під час щорічної вступної кампанії усі ВНЗ мусять дотримуватися визначених державою планів набору студентів за різними спеціальностями і напрямками. Оскільки зараз в Україні близько восьмисот діючих ВНЗ, постає проблема конкуренції ВНЗ за вступників. Для того, щоб підібрати ефективні заходи для залучення сильних абітурієнтів, кожному навчальному закладу необхідно виявити конкурентні ВНЗ, до яких вступник потенційно може подати оригінали документів під час зарахування.

За останні роки з'явилося багато джерел централізованої інформації, які можна використати для аналізу, та на його основі визначити конкурентні ВНЗ. Тому необхідно розробити механізм, який дає можливість сформулювати статистичні дані із відкритих джерел інформації та провести декілька експериментів, у результаті яких будуть виявлені конкурентні ВНЗ.

Мета і завдання

Метою магістерської дисертації є визначення конкурентів для НТУУ "КПІ", аналіз конкурентних переваг та надання рекомендацій щодо заходів по

залученню більшої кількості вступників. Для того, щоб досягнути поставленої мети, необхідно виконати наступний перелік завдань:

1. Дослідити умови та правила прийому до ВНЗ та визначити фактори, що впливають на вибір вступником ВНЗ.
2. Сформувавши порядок проведення експериментів, що дозволять визначити ВНЗ-конкуренти та дослідити відомі методи кластерного аналізу.
3. Сформувавши вимоги до системи, що буде формувати та завантажувати дані про вступників із обраного джерела інформації.
4. Розробити систему формування та завантаження даних у БД.
5. Виконати наведені експерименти та за отриманими даними назвати ВНЗ-конкуренти для НТУУ "КПІ", розробити рекомендації щодо заходів по залученню вступників до навчання в НТУУ "КПІ".

Рішення поставлених завдань та досягнуті результати

У даній роботі було досліджено переміщення абітурієнтів при поданні документів до ВНЗ. Проаналізовано фактори, які впливають на вибір вступником навчального закладу, у якому він буде отримувати вищу освіту та наведено порядок визначення конкурентів ВНЗ для НТУУ "КПІ". У роботі сформовано вимоги до бази даних, у якій будуть зберігатись дані про вступників, та модулю, що завантажує дані з відкритого джерела інформації про абітурієнтів ІС "Конкурс". Розроблену систему було протестовано. Результати тестувань показали, що при додаванні даних про вступників до бази було використано правильні критерії, перевірка яких дозволяє запобігти колізій у БД.

Також систему було протестовано на швидкодію. Створену БД було проаналізовано на предмет повноти даних. Дані тестування свідчать про те, що дані у БД повністю відповідають тим, що доступні на ІС "Конкурс".

За допомогою розробленої системи було проведено чотири експерименти. Отримані дані у результаті проведення цих експериментів дозволили визначити ВНЗ-конкуренти для НТУУ "КПІ". Навчальні заклади, які

можна назвати конкурентами НТУУ "КПІ" визначались у різних розрізах: за кількістю осіб, що подавали документи на усі можливі та окремо на технічні напрями, за кількістю заяв, та за сумою балів сертифікатів ЗНО і балу атестату про середню освіту. Також був проведений експеримент із застосуванням методів кластерного аналізу. За результатами цього експерименту можна зробити висновок, що вони підтвердили результати інших.

Об'єкт досліджень

Відповідно до поставленої мети об'єктом досліджень обрано щорічний прийом вступників до навчання у вищих навчальних закладах.

Предмет досліджень

Предметом досліджень у даній роботі обрано кореляцію між вищими навчальними закладами.

Методи досліджень

Методи статистичного та кластерного аналізу.

Наукова новизна

Наукова новизна роботи полягає у проведенні аналізу вступної кампанії та визначення ВНЗ-конкурентів НТУУ "КПІ" як для усіх відомих ВНЗ України так і для ВНЗ, що готують спеціалістів за тими напрямами, що є у НК "ПСА". Аналіз проведено із використанням методів статистичного аналізу та кластерного аналізу. За отриманими результатами аналізу розроблено рекомендації щодо оптимізації вступної кампанії для НТУУ "КПІ".

Практична цінність

Розроблена система може використовуватись будь-яким ВНЗ України для визначення конкурентів для себе. Також система та створена БД із даними про абітурієнтів може використовуватись для збору будь-яких статистичних

даних про результати вступної кампанії. На базі отриманих даних керівництво будь-якого ВНЗ може приймати керуючі рішення щодо стратегічного розвитку свого навчального закладу.

Ключові слова

Кластерний аналіз, статистичний аналіз, вступна кампанія, абітурієнт, ВНЗ, конкурент, база даних, ІС "Конкурс".

РЕФЕРАТ

на магистерскую диссертацию

выполненную на тему: Кластеризация перемещений абитуриентов при подаче документов в высшие учебные заведения

студентом: Патокой Егором Романовичем

Работа выполнена на 129 страницах, содержит 32 иллюстрации, 14 таблиц. При подготовке использовалась литература из 18 разных источников.

Актуальность

Одной из неотъемлемых частей жизненного цикла любого ВУЗа Украины является ежегодная вступительная кампания, целью которой является набор студентов на дневную, вечернюю или заочную форму обучения в соответствии с объемом государственного заказа, а также на обучение на контрактной основе.

Во время ежегодной вступительной кампании все ВУЗы должны придерживаться определенных государством планов набора студентов по различным специальностям и направлениям. Поскольку сейчас в Украине около восьмисот действующих ВУЗов, возникает проблема конкуренции ВУЗов за абитуриентов. Для того, чтобы подобрать эффективные меры для привлечения сильных абитуриентов, каждому учебному заведению необходимо выявить конкурентные ВУЗы, в которых поступающий потенциально может подать оригиналы документов при зачислении.

За последние годы появилось много источников централизованной информации, которые можно использовать для анализа, и на его основе определить конкурентные ВУЗы. Поэтому необходимо разработать механизм, который дает возможность сформировать статистические данные из открытых источников информации и провести несколько экспериментов, в результате которых будут выявлены конкурентные ВУЗы.

Цель и задачи

Целью магистерской диссертации является определение конкурентов для НТУУ "КПИ", анализ конкурентных преимуществ и рекомендации относительно мер по привлечению большего количества поступающих. Для того, чтобы достичь поставленной цели, необходимо выполнить следующий перечень задач:

1. Исследовать условия и правила приема в ВУЗы и определить факторы, влияющие на выбор поступающим ВУЗа.
2. Сформировать порядок проведения экспериментов, которые позволят определить ВУЗы-конкуренты и исследовать известные методы кластерного анализа.
3. Сформировать требования к системе, которая будет формировать и загружать данные о поступающих с выбранного источника информации.
4. Разработать систему формирования и загрузки данных в БД.
5. Выполнить указанные эксперименты и по полученным данным назвать ВУЗы-конкуренты для НТУУ "КПИ", разработать рекомендации по мерам по привлечению абитуриентов к обучению в НТУУ "КПИ".

Решение поставленных задач и достигнутых результатах

В данной работе были исследованы перемещения абитуриентов при подаче документов в ВУЗ. Проанализированы факторы, которые влияют на выбор поступающим учебного заведения, в котором он будет получать высшее образование и приведен порядок определения конкурентов для НТУУ "КПИ". В работе сформированы требования к базе данных, в которой будут храниться данные о поступающих, и модулю, который загружает данные из открытого источника информации об абитуриентах ИС «Конкурс». Разработанная система была протестирована. Результаты тестирования показали, что при добавлении данных о поступающих в базу были использованы правильные критерии, проверка которых позволяет предотвратить коллизий в БД.

Также система была протестирована на быстродействие. Созданная БД была проанализирована на предмет полноты данных. Данные тестирования свидетельствуют о том, что данные в БД полностью соответствуют тем, которые доступны на ИС "Конкурс".

С помощью разработанной системы было проведено четыре эксперимента. Полученные данные в результате проведения этих экспериментов позволили определить ВУЗы-конкуренты для НТУУ "КПИ". Учебные заведения, которые можно назвать конкурентами НТУУ "КПИ" определялись в различных разрезах: по количеству лиц, которые подавали документы на все возможные и отдельно на технические направления, по количеству заявлений, и по сумме баллов сертификатов ВНО и балла аттестата о среднем образовании. Также был проведен эксперимент с применением методов кластерного анализа. По результатам этого эксперимента можно сделать вывод, что они подтвердили результаты других.

Объект исследований

Согласно поставленной цели объектом исследований избран ежегодный прием абитуриентов на обучение в высших учебных заведениях.

Предмет исследований

Предметом исследований в данной работе избрана корреляция между высшими учебными заведениями.

Методы исследований

Методы статистического и кластерного анализа.

Научная новизна

Научная новизна работы заключается в проведении анализа вступительной кампании и определения ВУЗов-конкурентов НТУУ "КПИ" как для всех известных вузов Украины так и для ВУЗов, готовящих специалистов

по тем направлениям, которые есть в НК "ИПСА". Анализ проведен с использованием методов статистического анализа и кластерного анализа. По полученным результатам анализа разработаны рекомендации по оптимизации вступительной кампании для НТУУ "КПИ".

Практическая ценность

Разработанная система может использоваться любым ВУЗом Украины для определения конкурентов. Также система и создана БД с данными об абитуриентах может использоваться для сбора любых статистических данных о результатах вступительной кампании. На базе полученных данных руководство любого ВУЗа может принимать управленческие решения по стратегическому развитию своего учебного заведения.

Ключевые слова

Кластерный анализ, статистический анализ, вступительная кампания, абитуриент, ВУЗ, конкурент, база данных, ИС "Конкурс".

ABSTRACT

on master's thesis

on topic: Applicants Clustering during the Admission Campaign

Student: Yegor Patoka

Work carried out on 129 pages containing 32 images, 14 tables. The paper was written with references to 18 different sources.

Topicality

One of the integral parts of the life cycle of any university in Ukraine is annual introductory campaign, which aims to recruit students for full-time, evening or distance learning according to the volume of orders, as well as studying on a contract basis.

During the annual admission campaign all universities should follow the plans set by the state of students in different specialties and areas. Because now in Ukraine about eight existing universities, there is a problem of competition for university applicants. To pick up effective measures to attract strong applicants, each educational institution must identify competitive university, which potentially entrant may submit original documents during enrollment.

In recent years, many centralized sources of information that can be used for analysis, and on this basis to determine the competitive university. It is therefore necessary to develop a mechanism that makes it possible to generate statistical data from open sources of information and to conduct several experiments that result in competitive universities will be identified.

Aims and objectives

The purpose of the master's thesis is to determine the competitors for "KPI" to analyze competitive advantages and provide recommendations on measures to attract more enrollees. To achieve this goal, should be performed the following list of problems:

1. Investigate the conditions and rules of admission to universities and to identify factors that influence the choice of university entrant.
2. Form a procedure of experiments that will determine Universities competing and explore known methods of cluster analysis.
3. To form the system requirements that will form and upload the data on the source of the incoming information.
4. Develop a system for forming and loading data in the database.
5. Do the experiments and using the data obtained name universities that are competitors to "KPI". Create recommendations on measures to attract enrollees to study at NTUU "KPI".

The solution of the tasks and results

In this paper was investigated move students when submitting documents to the university. Factors that influence the choice of an enrollee of the institution in which he/she will receive a higher education and that are determining the order of competition for university "KPI". In this paper also determined requirements to the database in which to enrollee data will be stored, and to the module that loads data from open sources of information about enrollees IS "Konkurs". The developed system was tested. The results of tests showed that the data inserting criterias are correct and it helps to prevent collisions in the database.

Also, the system was tested on speed of performing. Established database was analyzed in terms of data completeness. These tests indicate that the data in the database fully correspond to those available in IS "Konkurs".

With help of the developed system four experiments were conducted. The data obtained as a result of these experiments allowed to identify competitors to "KPI". Educational institutions, which can be named as competitors to "KPI" determined in different sections: the number of people who filed documents on all possible and separately on technical directions, the number of applications and the number of mark got at exams and also avehigh school diploma. There was an

experiment with the use of cluster analysis. The results of this experiment confirmed the results of others.

The object of research

According to the aim object of study selected annual reception of enrollees for teaching in universities.

Subject of research

The subject of research in this paper selected correlation between higher education institutions.

Research Methods

Statistical methods and cluster analysis.

Scientific novelty

Scientific novelty lies in the opening campaign of the analysis and identification of universities competing "KPI" for all well-known universities in Ukraine and for universities that train specialists in those areas that are at ESC "IASA". The analysis was conducted using the methods of statistical analysis and cluster analysis. According to the results of analysis developed recommendations for optimizing the introductory campaign for "KPI".

The practical value of research

The system can be used in any university Ukraine to determine the competitors for themselves. The system and created a database with data about entrants can be used to collect any statistics on the results of the introductory campaign. Based on the data management of any university can take control of strategic decisions on their educational institution.

Keywords

Cluster analysis, statistical analysis, admission campaign, enrollee, university, competitor, database, IS "Konkurs".

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, СКОРОЧЕНЬ І ТЕРМІНІВ	18
ВСТУП.....	20
1 АНАЛІЗ ПРОЦЕСУ ВСТУПНОЇ КАМΠΑНІЇ.....	22
1.1 Аналіз процедури вступу до ВНЗ.....	22
1.2 Фактори, що впливають на обрання вступником ВНЗ, факультету(інституту) та напряму підготовки	24
1.3 Аналіз доступних джерел даних для дослідження	27
1.4 Висновки до першого розділу	29
2 МЕТОДИ ВИЗНАЧЕННЯ ПЕРЕМІЩЕНЬ АБІТУРІЄНТІВ ПРИ ПОДАЧІ ДОКУМЕНТІВ ДО ВНЗ	30
2.1 Визначення ВНЗ-конкурента.....	30
2.2 Інформаційна система "Конкурс"	32
2.3 Аналіз даних ІС "Конкурс"	39
2.4 Формування вимог до бази даних та модулю, що її формує	41
2.5 Основні засади кластерного аналізу.....	43
2.5 Висновки до другого розділу.....	48
3 РОЗРОБКА МОДУЛЯ ЗАВАНТАЖЕННЯ ТА ЗАПИСУ ДАНИХ З ІС "КОНКУРС" ДО БД.....	49
3.1 Формування бази даних рейтингів вступників.....	49
3.2 Завантаження даних абітурієнтів у базу	54
3.2.1 Краулер (CrawlerWrapper/Abot).....	57
3.2.2 Завантажувач (ConcurrentWebPageDownloader / WebClient)	58
3.2.3 Додання у БД (DatabaseWorker + БД)	59
3.2.4. Конвеєр	59
3.3 Висновки до третього розділу	60
4 РЕЗУЛЬТАТИ ПРОВЕДЕНИХ ЕКСПЕРИМЕНТІВ	61
4.1 Тестування роботи модуля.....	61

4.1.1	Перевірка правильності завантажених даних у БД	61
4.1.2	Оцінка продуктивності реалізованого модуля.....	65
4.2	Результати проведених експериментів	67
4.2.1	Експеримент 1: конкуренція ВНЗ за кількістю осіб	67
4.2.2	Експеримент 2: конкуренція ВНЗ за кількістю поданих заяв	73
4.2.3	Експеримент 3: Визначення ВНЗ-конкурента за допомогою методів кластерного аналізу.....	78
4.2.4	Експеримент 4: Визначення ВНЗ-конкурентів за сильних вступників	89
4.3	Висновки до четвертого розділу.....	97
5	ОХОРОНА ПРАЦІ ТА БЕЗПЕКА У НАДЗВИЧАЙНИХ СИТУАЦІЯХ	98
5.1	Вступ	98
5.2	Опис приміщення	98
5.3	Мікроклімат	101
5.4	Розрахунок системи кондиціонування	102
5.5	Шум.....	103
5.6	Випромінювання	105
5.7	Електробезпека	105
5.8	Пожежна безпека	106
5.9	Естетичні фактори виробничого середовища	108
5.10	Висновки до п'ятого розділу	108
	ВИСНОВКИ	110
	ПЕРЕЛІК ПОСИЛАНЬ	114
	ДОДАТОК А	117
	ДОДАТОК Б.....	118
	ДОДАТОК В.....	121

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, СКОРОЧЕНЬ І ТЕРМІНІВ

НТУУ "КПІ"	- Національний технічний університет України "Київський політехнічний інститут"
ПК	- Приймальна комісія
ЄДЕБО	- Єдина державна електронна база з питань освіти
ПІБ	- Прізвище ім'я побатькові
ЗНО	- Зовнішнє незалежне оцінювання
ВНЗ	- Вищий навчальний заклад
МОНУ	- Міністерство освіти і науки України
БД	- База даних
ОКР	- Освітньо-кваліфікаційний рівень
МАН	- Мала академія наук
ІС	- Інформаційна система
ЛНУ ім. Франка	- Львівський національний університет імені Івана Франка
НАУ	- Національний авіаційний університет
КНУ ім. Шевченка	- Київський національний університет імені Тараса Шевченка
НУ "ЛП"	- Національний університет "Львівська політехніка"
КНЕУ ім. Гетьмана	- Київський національний економічний університет імені Вадима Гетьмана, ДВНЗ
КНТЕУ	- Київський національний торговельно-економічний університет
НУХТ	- Національний університет харчових

		технологій
НПУ	ім.	- Національний педагогічний університет імені М.П. Драгоманова
Драгоманова		
ХНУ	ім.	- Харківський національний університет імені В.Н. Каразіна
Каразіна		
ЧНУ	ім.	- Чернівецький національний університет імені Юрія Федьковича
Федьковича		
ОНУ	ім.	- Одеський національний університет імені І.І. Мечникова
Мечникова		
НУКМА		- Національний університет "Києво-Могилянська академія"
НТУ "ХПІ"		- Національний технічний університет "Харківський політехнічний інститут"
ВНТУ		- Вінницький національний технічний університет
ХНУРЕ		- Харківський національний університет радіоелектроніки
ОНПУ		- Одеський національний політехнічний університет
ЗНТУ		- Запорізький національний технічний університет
НАУ ХАІ		- Національний аерокосмічний університет ім. М.Є. Жуковського "Харківський авіаційний інститут"
ДВНЗ НГУ		- Державний вищий навчальний заклад "Національний гірничий університет"
КНУБА		- Київський національний університет будівництва та архітектури
ІТ		- Інформаційні технології

ВСТУП

Однією з невід'ємних частин життєвого циклу будь-якого ВНЗ України є щорічна вступна кампанія, метою якої є набір студентів на денну, вечірню або заочну форму навчання відповідно до обсягу державного замовлення, а також на навчання на контрактній основі.

Під час щорічної вступної кампанії усі ВНЗ мусять дотримуватися визначених державою планів набору студентів за різними спеціальностями і напрямками. Цьому приділяється значна увага, адже, у разі недобору абітурієнтів, МОНУ може застосувати до освітніх установ санкції, як то скорочення кількості бюджетних місць або, у крайньому випадку, т.зв. оптимізація – об'єднання ВНЗ.

Разом з цим, кожен ВНЗ зацікавлений у наборі найбільш сильних абітурієнтів, тому що якість проведеної вступної кампанії також впливає на рейтинг ВНЗ серед роботодавців, адже вони приймають рішення про зарахування на роботу випускників на основі рівня їхніх знань.

Для того щоб ВНЗ міг надати високий рівень знань своїм студентам, необхідно набрати якнайбільше тих вступників, які добре підготовлені до навчання у цьому ВНЗ. Тому навчальний заклад має вжити необхідних заходів перед вступною кампанією, щоб залучити якомога більше сильних абітурієнтів. Такими заходами можуть бути різноманітні види рекламних кампаній, співпраця зі школами, створення факультетів доуніверситетської підготовки та багато інших.

Оскільки зараз в Україні близько восьмисот діючих ВНЗ, постає проблема конкуренції ВНЗ за вступників. Для того, щоб підібрати ефективні заходи для залучення сильних абітурієнтів, кожному навчальному закладу необхідно виявити конкурентні ВНЗ, до яких вступник потенційно може подати оригінали документів під час зарахування.

За останні роки з'явилося багато джерел централізованої інформації, які можна використати для аналізу, та на його основі визначити конкурентні ВНЗ. Тому необхідно розробити механізм, який дає можливість сформувати статистичні дані із відкритих джерел інформації та провести декілька експериментів, у результаті яких будуть виявлені конкурентні ВНЗ.

1 АНАЛІЗ ПРОЦЕСУ ВСТУПНОЇ КАМΠΑНІЇ

1.1 Аналіз процедури вступу до ВНЗ

Щороку МОНУ оприлюднює умови прийому до ВНЗ у вигляді законодавчого документу – "Умови прийому на навчання до вищих навчальних закладів України". Ці умови є обов'язковими для вищих навчальних закладів незалежно від форми власності та підпорядкування. Організацію прийому вступників до вищого навчального закладу здійснює приймальна комісія (відбіркова комісія – для вищих навчальних закладів I-II рівнів акредитації), склад якої затверджується наказом керівника вищого навчального закладу, який є її головою.

Кожен ВНЗ до початку вступної кампанії має розробити, затвердити и оприлюднити правила прийому до ВНЗ. Правила розробляються відповідно до чинного законодавства України, у тому числі й умов прийому, і затверджуються вченою радою ВНЗ. У переліку обов'язкових правил має бути також зазначено[1]:

- перелік напрямів підготовки (спеціальностей), за якими вищий навчальний заклад оголошує прийом на навчання, із зазначенням факультетів або інститутів), на яких здійснюється підготовка за кожним напрямом (спеціальністю);

- кількість місць ліцензованого обсягу за кожним напрямом підготовки (спеціальністю) для кожного освітньо-кваліфікаційного рівня з розподілом по структурних підрозділах (факультетах, інститутах тощо), на яких здійснюється підготовка;

- порядок проведення конкурсного відбору та строки зарахування вступників за різними ОКР;

- перелік конкурсних предметів із зазначенням профільних, з яких вступники подаватимуть сертифікат ЗНО для вступу на основі повної загальної середньої освіти;
- порядок нарахування додаткових балів особам, яким це право надано;
- порядок обчислення конкурсного бала вступника;
- значення мінімальної кількості балів з кожного предмета у сертифікаті зовнішнього незалежного оцінювання (вступних екзаменів), з якими вступник може бути допущений до участі в конкурсі або зарахований на навчання поза конкурсом, але не нижче 124 балів з непрофільних та профільних предметів.

У період вступної кампанії вступники подають заяву на участь у конкурсному відборі до ВНЗ в паперовій або в електронній формі. Заява в паперовій формі подається вступником особисто до ПК навчального закладу.

При вступі на навчання за різними ОКР абітурієнт може подати заяву не більше ніж до п'яти ВНЗ України та не більше ніж на три напрями підготовки (на три спеціальності) у кожному з них. Заяви, подані на певний напрям підготовки (спеціальність) до одного ВНЗ за різними формами навчання, вважаються фактом подання однієї заяви. Факт подання заяви фіксується у ЄДЕБО.

Під час подання заяв про участь у конкурсному відборі вступник зазначає у кожній заяві пріоритетність цієї заяви відносно інших поданих ним заяв, при цьому показник пріоритетності 1 означає найвищу пріоритетність.

У випадку, коли навчальний заклад здійснює підготовку за певним напрямом на різних факультетах (інститутах), вступник зазначає пріоритетність зарахування за обраним напрямом відповідно до факультету (інституту).

У випадку, коли у ВНЗ підготовка за напрямом здійснюється за спрямуваннями (спеціалізацією, нозологією, профілем, мовою тощо), вступник зазначає у кожній заяві обрану ним пріоритетність зарахування за відповідним спрямуванням.

Після оприлюднення списків рекомендованих до зарахування ПК, вступник у випадку, якщо він бажає вступити має особисто подати оригінали документів. Якщо вступник не зробив цього у встановлений термін, то він втрачає право на зарахування на навчання.

Якщо абітурієнта рекомендовано до зарахування на навчання у ВНЗ, який було позначено найвищим пріоритетом, абітурієнт втрачає право брати участь у конкурсному відборі на навчання у ВНЗ з нижчим пріоритетом. За рішенням ПК заяви з нижчим пріоритетом скасовуються в рейтингових списках, сформованих в ЄДЕБО. Якщо вступника рекомендовано до зарахування на навчання до другого за пріоритетом ВНЗ або нижче, абітурієнт втрачає право на участь у подальшому конкурсному відборі на навчання за наступними пріоритетами, при цьому зберігається його право на участь у конкурсі за вищими пріоритетами. За абітурієнтом залишається право на вступ до ВНЗ, де його заяву було скасовано, за рахунок фізичних або юридичних осіб [1].

1.2 Фактори, що впливають на обрання вступником ВНЗ, факультету(інституту) та напряму підготовки

У попередньому пункті подано витримки із офіційних умов прийому до ВНЗ для 2014 року[1]. Якщо проаналізувати умови вступу теперішнього, минулого та попередніх років, то можна виявити в них відмінності, подекуди – досить суттєві. У зв'язку з цим, умови вступу, які будуть безпосередньо впливати на вибір вступником ВНЗ, факультету (інституту), також змінюватимуться щороку. Принаймні, з огляду на досвід попередніх років, така тенденція зберігатиметься у найближчі роки. Тому, для того щоб результати аналізу були максимально наближені до тих, які буде отримано наприкінці вступної кампанії 2015 року, необхідно аналізувати дані останньої вступної кампанії і спиратися на ті умови, що діяли під час її проведення. Тому у якості об'єкту дослідження було обрано вступну кампанію 2014 року.

Отже, за умов вступу, що діяли у 2014 році[1], можна виокремити такі фактори, що впливали на вибір абітурієнтом ВНЗ, факультету (інституту):

1) Кількість ВНЗ, до яких вони можуть вступити, – не більше п'яти, а кількість напрямів, за якими вступник має право подати заяву на навчання, – не більше трьох. Якщо раніше вступник міг обрати будь-який ВНЗ із восьми сот (як, наприклад, у 2009-2010 рр.) і подати заяви хоч у кожен із них, то наразі він такої можливості не має[1].

2) Порядок підрахунку бонусних балів кожен ВНЗ визначає по своєму, згідно з власними правилами прийому. Так, наприклад, деякі ВНЗ можуть мати курси доуніверситетської підготовки, факт проходження яких враховується при вступі до ВНЗ.

3) Обсяг державного замовлення. Навіть якщо конкурсний бал у вступника досить високий, через невелику кількість бюджетних місць абітурієнт може не вступити до бажаного ВНЗ, і тому подати оригінали до іншого.

4) Згідно з умовами прийому 2015 року, особа втрачає право на вступ до менш пріоритетних ВНЗ у випадку, коли її рекомендовано до більш пріоритетного ВНЗ.

Наведені вище фактори можна використовувати при визначенні сильних сторін ВНЗ-конкурентів, а також при формуванні подальших рекомендацій щодо заходів до залучення більшої кількості сильних вступників.

Окрім факторів, що зумовлені умовами вступу, також є й інші, більш суб'єктивні фактори, що впливають на вибір абітурієнта, особливо на етапі формування списку ВНЗ, розставлених за пріоритетами. Наприклад, близькість ВНЗ до місця проживання вступника, адже абітурієнт може не мати змоги навчатись у більш престижному, проте віддаленому ВНЗ.

На рішення абітурієнтів щодо вступу можуть також впливати різноманітні рейтингові списки ВНЗ, що розробляються аналітичними або іншими організаціями. Серед відомих рейтингів ВНЗ можна виділити такі:

- Рейтинг українських ВНЗ «Компас»;
- QS World University Rankings;
- Рейтинг "Топ-200 Україна".

Мета рейтингу українських ВНЗ "Компас", за словами його творців, це створення переліку ВНЗ України з їх об'єктивною та прозорою оцінкою за такими параметрами:

- практична цінність послуг вищої освіти для студентів;
- відповідність освітніх послуг вимогам реального сектора економіки;
- надання випускникам надійних перспектив працевлаштування після закінчення навчання.

Рейтинг ВНЗ України «Компас» створюється на основі спеціально розробленої методології, яка враховує дані, отримані в результаті комплексу соціологічних досліджень:

- опитування компаній-роботодавців;
- опитування випускників 2007-2012 років, які закінчили ВНЗ, що беруть участь у рейтингу;
- опитування експертів.

У формуванні рейтингу взяло участь 239 ВНЗ, відібраних за критеріями відбору творців рейтингу "Компас"[3].

Популярними серед українських абітурієнтів є також міжнародні рейтинги, як, наприклад, QS World University Rankings. Рейтинг складається за чотирма найбільш важливими для потенційних студентів критеріями: дослідження, викладання, працевлаштування та міжнародні перспективи. Ці ключові сфери оцінюються за допомогою шести показників, кожен з яких має свій максимальний ваговий коефіцієнт у відсотках. Чотири з них засновані на об'єктивних даних, ще два ґрунтуються на матеріалах глобальних опитувань,

учасниками одного з яких є найшанованіші в своїй царині науковці, другого – найавторитетніші в своїй галузі роботодавці.

Також відомий в Україні і в Європі є рейтинг "Топ-200 Україна"[9]. Незмінним базисом методики залишався загальний набір індикаторів для університетів різних типів, що дозволяє порівнювати результати їхньої діяльності. Діяльність вузу оцінюється з допомогою інтегрованого індексу — Із. Цей індекс містить три комплексні складові: $I_z = I_{np} + I_n + I_{mv}$, де I_{np} — індекси якості науково-педагогічного потенціалу, I_n — індекс якості навчання, I_{mv} — індекс міжнародного визнання. Методику створено відповідно до Берлінських принципів ранжування університетів[9].

Усього для формування зазначених індексів нинішнього року використовувалися 24 індикатори прямого виміру з сумарною вагою 80% і експертне оцінювання з ваговим коефіцієнтом 20%. Цього року для оцінювання міжнародного визнання вузу додатково враховувалася його участь у таких європейських програмах, як Seventh Framework Programme і Tempus. Дані про ці індикатори, за результатами діяльності університетів у попередньому році, вибиралися з кількох незалежних джерел. Експертне оцінювання проводилося за такими критеріями: рівень базової, загальноосвітньої підготовки студентів, рівень їхньої фахової підготовки, рівень практичного володіння інформаційними технологіями, затребуваність випускників вузів ринком праці.

1.3 Аналіз доступних джерел даних для дослідження

Для того щоб сформувати статистику, на базі якої можна визначити ВНЗ-конкуренти для НТУУ "КПІ", необхідно обрати джерела інформації, звідки можна буде отримати необхідні дані. Першоджерелом даних про заяви вступників є Єдина державна електронна база з питань освіти (ЄДЕБО).

ЄДЕБО – автоматизована система накопичення, оброблення, зберігання та захисту даних, у тому числі персональних, які використовуються закладами, що надають освітні послуги в Україні. Власником (утримувачем) ЄДЕБО є

МОНУ, адміністратором (розпорядником) – державне підприємство «Інфоресурс», що належить до сфери управління МОНУ.

При прийомі документів та заяв від абітурієнтів оператори та адміністратор ЄДЕБО від навчального закладу мають вносити ці дані у базу[3]. Процедура занесення даних відбувається у відповідності до [4]. У ЄДЕБО зберігаються дані про учнів шкіл, абітурієнтів та студентів, а отже й інформація про подані абітурієнтами заяв також є в ЄДЕБО. Але доступ до бази обмежений, тому за допомогою ЄДЕБО отримати дані про вступників неможливо.

З метою оперативного інформування абітурієнтів про кількість поданих заяв на навчання за освітньо-професійною програмою бакалавра (спеціаліста, магістра медичного та ветеринарно-медичного спрямувань) відповідних напрямів (спеціальностей) у вищих навчальних закладах III-IV рівнів акредитації було запроваджено інформаційну систему "Конкурс. Дані до ІС потрапляють безпосередньо з ЄДЕБО[5].

ІС "Конкурс" має рейтингові списки вступників до кожного з існуючих в Україні ВНЗ, усіх факультетів(інститутів), усіх спеціальностей. Іншим відкритим джерелом інформації, доступним у мережі Інтернет, є "Сервіс пошуку абітурієнтів Абіт-пошук". У Абіт-пошук дані потрапляють з ІС "Конкурс". Цей ресурс також має рейтингові списки абітурієнтів. На сайті реалізована можливість пошуку вступника за його ПІБ та балом атестата. До того ж, ресурс має рейтинги ВНЗ за різними показниками, наприклад, за кількістю поданих заяв, середнім прохідним конкурсним балом тощо.

Для аналізу більш доцільно використовувати дані з ІС "Конкурс", оскільки інформація, що доступна на цьому ресурсі отримується безпосередньо з ЄДЕБО. Абіт-пошук отримує дані із ІС "Конкурс", а отже потенційно при завантаженні даних із ІС "Конкурс" до Абіт-пошуку можуть не потрапляти деякі дані.

Оскільки у роботі будуть використовуватись персональні дані про особу, то виникає необхідність упевнитись, що їх використання не буде порушувати чинне законодавство України щодо захисту персональних даних.

Згідно з [6], персональні дані не можуть збиратись та оброблюватись без попередньої згоди власника персональних даних, а отже, якщо у системі "Конкурс" та ЄДЕБО розміщено інформацію про абітурієнта, то це означає, що ця особа попередньо вже дала згоду на збір та обробку її персональних даних, занесення їх до ЄДЕБО, а також публікацію деяких даних в мережі Інтернет. Згідно цього ж закону, персональні дані, крім знеособлених персональних даних, є інформацією з обмеженим доступом. Знеособлення персональних даних передбачає вилучення відомостей, які дають змогу повністю чи частково ідентифікувати особу. Дана підсистема не буде використовувати персональні дані, які можуть прямо ідентифікувати особу, а саме серійні номери та номери документів про середню освіту, сертифікатів ЗНО, паспорту, водійських посвідчень, ідентифікаційного коду та інші дані, що не є знеособленими.

1.4 Висновки до першого розділу

У даному розділі проведено аналіз процесу вступної кампанії. Розглянуто фактори, зумовлені правилами вступу до ВНЗ, а також наведено приклад інших, суб'єктивних факторів, що впливають на вибір вступником того чи іншого ВНЗ. Проаналізовано можливі альтернативні джерела додаткової інформації про вступника та обрано джерело, дані з якого будуть використовуватись для дослідження. Також у розділі наведено аналіз закону "Про захист персональних даних".

2 МЕТОДИ ВИЗНАЧЕННЯ ПЕРЕМІЩЕНЬ АБІТУРІЄНТІВ ПРИ ПОДАЧІ ДОКУМЕНТІВ ДО ВНЗ

2.1 Визначення ВНЗ-конкурента

Перед тим, як почати дослідження щодо визначення конкуруючих ВНЗ, необхідно визначитись із тим, який ВНЗ може вважатись конкуруючим. Як вже зазначалося, кожен ВНЗ змагається з іншими ВНЗ за кількість вступників. А отже, їм необхідно, щоб до ВНЗ було подано якомога більше заяв на вступ. Нехай маємо два навчальних заклади, у яких приблизно однакова кількість поданих заяв, а перелік напрямів підготовки та спеціальностей буде однаковий. За таких умов, ці два ВНЗ будуть явними конкурентами, причому, переможцем буде той ВНЗ, у якого кількість студентів першого курсу після завершення вступної кампанії буде більшою.

У даній схемі є вузьке місце: серед усіх ВНЗ України майже немає ВНЗ, у яких би співпадав перелік бакалавратів. А отже, виявити прямих конкурентів неможливо. З іншого боку, оскільки перелік усіх спеціальностей досить великий, то вірогідним є випадок, коли у одному ВНЗ є напрями підготовки, подібні до напрямів підготовки у іншому закладі, тобто відмінності у навчальній програмі незначні. Тому, маючи вибір між двома такими ВНЗ, абітурієнт має прийняти рішення.

У такому випадку є сенс порівнювати ВНЗ не за кількістю поданих заяв, а за кількістю унікальних вступників, що подали документи на вступ до двох ВНЗ одночасно. Такий експеримент необхідно провести для кожного із ВНЗ. Це дасть змогу встановити, за яку кількість майбутніх студентів конкурують навчальні заклади.

Кількості вступників, які одночасно подали документи до двох ВНЗ одночасно, буде недостатньо для того, щоб зрозуміти, наскільки сильним є конкурентний ВНЗ. Для того щоб зробити таку оцінку, необхідно визначити відсоток спільник абітурієнтів до загальної кількості абітурієнтів, що подали

документи на вступ. Щоб зрозуміти, який з двох ВНЗ отримав перевагу, необхідно визначити кількість вступників, що подали оригінали документів до одного із ВНЗ, що розглядаються.

У даній роботі буде визначено конкурентів НТУУ "КПІ". Оскільки НТУУ "КПІ" є технічним навчальним закладом, то необхідно, окрім загального експерименту (тобто врахування загальної кількості вступників), провести аналогічний експеримент, але тільки серед вступників, що подавали документи за технічними напрямами. У даній роботі, у інтересах кафедри Системного проектування, а також Інституту прикладного системного аналізу НТУУ "КПІ" у якості експериментальних напрямів будуть обрані наступні:

- Комп'ютерні науки;
- Програмна інженерія;
- Системна інженерія;
- Системний аналіз;
- Інформатика.

Ці напрями належать до галузей "Інформатика та обчислювана техніка" і "Системні науки та кібернетика". Як раніше зазначалось, в інших галузях є такі напрями як, наприклад, "Комп'ютерна інженерія", що є досить схожою за характером підготовки студентів. А отже, потенційно, ВНЗ, що готує спеціалістів за таким напрямом, також може бути конкурентом для НТУУ "КПІ". Але у даній роботі будуть розглядатись тільки вище перераховані напрямки.

Таким чином, виконавши експерименти, що описані вище, можна отримати інформацію про ті ВНЗ, що є конкурентами НТУУ "КПІ" та який із них отримав перевагу.

У першому розділі зазначалось, що окрім того, що ВНЗ змагаються за кількість вступників, конкуренція є і за сильних абітурієнтів, тобто тих, які добре підготовлені до навчання у тому чи іншому ВНЗ. Отже, виникає необхідність визначити конкуренти ВНЗ ще й у цьому розрізі. Найкращим показником підготовленості абітурієнта до навчання, з точки зору вищого

навчального закладу є сума сумарного балу ЗНО та балу атестата про середню освіту, тому необхідно провести експеримент серед тих осіб, що вступили до ВНЗ. Таких осіб необхідно згрупувати так, щоб можна було визначити, до якого ВНЗ найчастіше вступають абітурієнти із певною сумою балу ЗНО та балу атестата про середню освіту. Для проведення такого експерименту необхідно використовувати методи кластерного аналізу.

Метою даної роботи також є надання рекомендацій НТУУ "КПІ" щодо проведення заходів для залучення більш сильних вступників до навчання. Тому, визначивши конкурентів ВНЗ, необхідно також виявити сильні сторони ВНЗ у порівнянні із НТУУ "КПІ". Щоб виявити сильні сторони ВНЗ-конкурентів, зумовлені умовами вступу до ВНЗ, наприклад такими, як середній прохідний бал, обсяг бюджетних місць, врахування різних додаткових балів та інші, необхідно провести певні експерименти. Перше, що необхідно для проведення цих експериментів і надання рекомендацій – це джерело інформації, звідки можна отримати дані для подальшого аналізу. У якості такого джерела використовуємо використовувати ІС "Конкурс".

2.2 Інформаційна система "Конкурс"

Як згадувалось раніше, ІС "Конкурс" було створено задля оперативного інформування абітурієнтів про кількість поданих заяв на навчання за освітньо-професійною програмою бакалавра (спеціаліста, магістра медичного та ветеринарно-медичного спрямувань) відповідних напрямів (спеціальностей) у вищих навчальних закладах III-IV рівнів акредитації.

Доступ до даного ресурсу є відкритим, тобто кожен бажаючий має змогу отримати інформацію із сайту. Для того, щоб побачити контент ІС, достатньо використати будь-який з відомих веб-браузерів та у адресному полі ввести: "http://vstup.info". Головна сторінка ІС відображена на рис. 2.1:

Інформаційна система «Конкурс»

Вступ до вищих навчальних закладів України I-IV рівнів акредитації

Вступна кампанія 2014

Сторінку оновлено 27/09/14 о 22:52 за даними ВНЗ з допомогою Єдиної державної електронної бази з питань освіти.

**Умови прийому до вищих навчальних закладів України в 2014 році (на сайті МОН України)
Функціонування та умови використання ІС «Конкурс» в 2014 році**

Статистика

	Всього	Мол. спец.	Бакалавр	Спеціаліст	Магістр
Подано заяв:	2125846	329797	1392395	226890	176764
Допущено до конкурсу:	1663712	211512	1121764	189377	141059
Серед допущених: заяв із правом позаконкурсного вступу:	37327	8511	24903	2130	1783
Серед допущених: подано оригіналів документів:	662513	148573	261634	151234	101072
Серед допущених: рекомендовано до зарахування:	155533	17471	128126	5815	4121
Серед допущених: зараховано:	588095	132002	233666	135852	86575
Ліцензований обсяг:	2358030	399628	1267751	489302	201349
Бюджетних місць:	315529	96214	117868	59436	42011
Загальний конкурс, заяв на місце:	0.71	0.53	0.88	0.39	0.7
Конкурс на бюджетні місця, заяв на місце:	5.27	2.2	9.52	3.19	3.36

Пошук абітурієнтів

Пошук заяв абітурієнтів за прізвищем, іменем та по-батькові в межах поточної вступної кампанії

Рисунок 2.1 - Головна сторінка ІС "Конкурс"

У верхній частині ресурсу зображена статистика за останній рік вступної кампанії (у даному випадку – 2014). Також зазначено, що дані оновлюються згідно з ЄДЕБО.

Отже, як видно з зображення, для того щоб перейти до рейтингового списку, необхідно пройти п'ять рівнів посилань. Для початку, треба обрати рік вступної кампанії. За замовчуванням обирається поточний або останній рік вступної кампанії. Далі необхідно обрати, за яким критерієм шукати ВНЗ. ІС "Конкурс" пропонує здійснювати пошук за наступними критеріями (рис.2.2):

- за алфавітом;
- за тегами;
- за галуззю знань;
- за територією.



Рисунок 2.2 - Пошук ВНЗ за різними критеріями

Після обрання критерію пошуку ВНЗ буде запропоновано список ВНЗ, що відповідають обраному критерію. Наприклад, якщо обрати територіальний пошук, м. Київ, то в наступній сторінці відкриється список усіх ВНЗ, що відповідають цьому критерію (рис.2.3).

Вступна кампанія 2012 » Пошук за регіонами: м. Київ

Сторінку оновлено 29/08/12 о 17:20 за даними Єдиної державної електронної бази з питань освіти.

м. Київ

Академія муніципального управління

Академія праці і соціальних відносин Федерації професійних спілок України

Вищий навчальний заклад "Київський медичний коледж ім.П.І.Гаврося"

Вищий навчальний заклад "Київський університет ринкових відносин"

Вищий навчальний заклад "Коледж театру і кіно"

Вищий навчальний заклад "Національна академія управління"

Вищий навчальний заклад "Перший Київський медичний коледж"

Вищий навчальний заклад "Університет економіки та права "КРОК"

Вищий навчальний заклад Київський медичний коледж №3

Вищий навчальний заклад Київський міський медичний коледж

Відкрите акціонерне товариство Вищий навчальний заклад "Київський міжгалузевий інститут підвищення кваліфікації"

Відкритий міжнародний університет розвитку людини "Україна"

Відокремлений структурний підрозділ "Київський педагогічний коледж східних мов" при ПВНЗ "Київський університет "Східний світ"

ВСП "Інститут інтелектуальної власності Національного університету "Одеська юридична академія" в м. Києві"

Державна установа "Інститут економіки та прогнозування Національної академії наук України"

Державний вищий навчальний заклад "Київський енергетичний коледж"

Державний вищий навчальний заклад "Київський індустріальний коледж"

Рисунок 2.3 - ВНЗ у місті Києві

Далі відкривається "профайл" обраного ВНЗ із загальною інформацією про ВНЗ, тобто контактними даними, адресою, статистикою та посиланнями на форми навчання, а саме (рис. 2.4):

- денна;
- заочна.

Вступна кампанія 2012 » Національний технічний університет України «Київський політехнічний інститут»

Сторінку оновлено 29/08/12 о 17:21 за даними Єдиної державної електронної бази з питань освіти.

Про ВНЗ	Статистика	Денна форма	Заочна форма
Назва ВНЗ:	Національний технічний університет України «Київський політехнічний інститут»		
Тип ВНЗ:	вищий навчальний заклад III-IV рівнів акредитації		
Поштовий індекс:	03056		
Адреса:	М.КИЇВ, вул. проспект Перемоги, 37,		
Телефони:	+ 38 (044)454-96-45, 454-96-46, 406-85-57		
Веб-сайт:	www.ntu-kpi.kiev.ua		
E-mail:	post [at] ntu-kpi.kiev.ua		

Рисунок 2.4 - "Профайл" ВНЗ

Наприклад, якщо обрати денну форму навчання, то на наступній сторінці відкриється список пропозицій освітньо-кваліфікаційних рівнів усіх можливих галузей та напрямків в обраному раніше ВНЗ (рис. 2.5).

Перелік пропозицій денної форми навчання:

ОКР	Н	Σ	Л	Д	П
Бакалавр (повний термін навчання)	Факультет: Факультет біотехнології і біотехніки Галузь: Машинобудування та матеріалобробка Напрямок: машинобудування	конкурс	30	20	Математика (профільний) Українська мова та література Фізика (профільний)
Бакалавр (повний термін навчання)	Факультет: Механіко-машинобудівний інститут Галузь: Машинобудування та матеріалобробка Напрямок: машинобудування	конкурс	135	50	Математика (профільний) Українська мова та література Фізика (профільний)
Бакалавр (повний термін навчання)	Факультет: Факультет електроенерготехніки та автоматики Галузь: Електротехніка та електромеханіка Напрямок: електромеханіка	конкурс	165	67	Математика (профільний) Українська мова та література Фізика (профільний)
Бакалавр (повний термін навчання)	Факультет: Механіко-машинобудівний інститут Галузь: Машинобудування та матеріалобробка Напрямок: інженерна механіка	конкурс	270	136	Математика (профільний) Українська мова та література

Рисунок 2.5 - Перелік пропозицій НТУУ "КПІ" денної форми навчання

Потім, після обрання бажаного факультету обраного ВНЗ та спеціальності, при натисненні на посилання "Конкурс" відкривається список абітурієнтів, що подали заяви на вступ на обрану користувачем спеціальність, факультет, ВНЗ. Приклад одного із списків зображено на рис. 2.6.

**Вступна кампанія 2012 » Національний технічний університет України
«Київський політехнічний інститут» » » Список абітурієнтів на денній формі
навчання**

Сторінку оновлено 29/08/12 о 18:14 за даними Єдиної державної електронної бази з питань освіти.

#	ПІБ	Σ	ЗНО	Е	С	О	П	ПК	ПЧ	Д
1	Андрук Микола Миколайович	693.7	508.00	0.00	185.70	—	—	+	—	+
2	Дух Дмитро Васильович	747.4	545.50	0.00	182.90	—	19.00	—	—	+
3	Сиротін Олександр Васильович	724.8	530.50	0.00	174.30	—	20.00	—	—	+
4	Куліков Дмитро Геннадійович	720.4	533.00	0.00	167.70	—	19.70	—	—	+
5	Періжок Денис Дмитрович	714.9	522.50	0.00	192.40	—	—	—	+	+
6	Царик Євгеній Вікторович	711.6	533.50	0.00	178.10	—	—	—	—	+
7	Яковець Назар Сергійович	709.5	508.50	0.00	181.00	—	20.00	—	—	+
8	Піонткевич Іван Олександрович	707.6	520.00	0.00	187.60	—	—	—	+	+
9	Сердюк Артем Сергійович	707.6	529.50	0.00	178.10	—	—	—	—	+

Рисунок 2.6 - Приклад списку абітурієнтів, що подали зави на вступ.

Отже, як видно з рисунків вище у системі, рейтинг абітурієнтів подано у досить неявному вигляді. Необхідно виконати п'ять кроків перш ніж побачити список абітурієнтів, які подали документи на конкретну спеціальність, факультет та ВНЗ. Також з рисунків видно, що прямого пошуку інформації про особу немає, тобто, використовуючи функціонал веб-сайту, неможливо дізнатись, куди ще подала заяву на вступ на перший курс та чи інша особа. Загальна структура веб-ресурсу наведена у "Додатку А".

2.3 Аналіз даних ІС "Конкурс"

Згідно з [5], ІС "Конкурс" містить у собі інформацію про рейтинги абітурієнтів, що подали зави на вступ у ВНЗ України усіх можливих рівнів акредитації (I - IV), усіх можливих ОКР.

Рейтинг абітурієнтів в обраному користувачем ВНЗ, факультеті або інституті, напрямку та галузі освіти та ОКР подається у вигляді таблиці із одинадцятьма стовпчиками:

- "#" – номер особи у рейтингу;
- "ПБ" – прізвище, ім'я та по батькові абітурієнта;
- "Σ" – сума балів сертифікатів ЗНО, що обов'язкові для обраної спеціальності, в обраному ВНЗ та середнього балу атестата про середню освіту
- "ЗНО" – сума балів сертифікатів ЗНО, що обов'язкові для обраної спеціальності;
- "Е" – бали вступних іспитів (якщо такі відбувались);
- "С" – середній бал атестата про середню освіту абітурієнта;
- "О" – додаткові бали за призові місця у всеукраїнських шкільних олімпіадах та конкурсів МАН з предметів, що є обов'язковими для обраної спеціальності в обраному ВНЗ;
- "П" – бали за курси доуніверситетської підготовки;
- "ПК" – право на позаконкурсний вступ;
- "ПЧ" – право на позачерговий вступ;
- "Д" – наявність оригіналів документів про середню освіту та сертифікатів ЗНО.

Кількість рядків залежить від кількості осіб, що подали документи та заяву на вступ на перший курс.

Таблиця є HTML-сторінкою та зроблена за допомогою HTML-тегів, які відповідають за створення та редагування таблиць. Дані в таблиці є строковими.

Важливою вимогою до джерела інформації є наявність даних у рейтингах про абітурієнтів, що мали прохідний бал на момент вступної кампанії, але не подали оригіналів документів, оскільки за їх відсутності, буде неможливо визначити точну кількість спільних вступників. Проаналізувавши ІС "Конкурс" видно, що система надає такі дані (див. рис. 2.7):

84	Бабич Дмитро Андрійович	515.2	40.5	Фізика: 158.5 Укр.м.: 163.0 Матем.: 174.0	—	-/19.7	—	+	—	+	
85	Хмара Олександр Ігорович	511.5	46.0	Фізика: 181.5 Матем.: 165.5 Укр.м.: 157.5	—	-/-	—	—	—	+	
86	Зворський Олексій Олексійович	510.0	46.0	Фізика: 165.5 Матем.: 174.0 Укр.м.: 166.0	—	-/-	—	—	—	+	
87	Ванеєв Іван Олексійович	497.5	51.5	—	Матем.: 168.0 Укр.м.: 136.0 Фізика: 176.0	-/-	—	—	—	+	
88	Диренков Владислав Олексійович	485.4	40.0	Фізика: 160.5 Матем.: 167.5 Укр.м.: 132.0	—	-/18.4	—	+	—	+	
89	Шелупець Максим Валерійович	476.0	41.0	Фізика: 146.0 Матем.: 152.5 Укр.м.: 155.0	—	-/20.0	—	+	—	+	
90	Осінній Дмитро Віталійович	469.5	51.0	Фізика: 156.0 Матем.: 149.0 Укр.м.: 151.0	—	-/-	—	—	—	+	
91	Мельничук Володимир Олегович	528.5	52.5	Фізика: 168.5 Матем.: 180.0 Укр.м.: 170.0	—	-/-	+	+	—	-/+	3/0
92	Іванков Юрій Геннадійович	524.5	54.0	Фізика: 168.5 Укр.м.: 155.5 Матем.: 185.0	—	-/-	+	+	—	-/+	11/0
93	Московенко Євгеній Олегович	522.5	44.0	Фізика: 176.0 Матем.: 172.0 Укр.м.: 174.0	—	-/-	+	—	—	-/+	4/0
94	Штода Дмитро Олександрович	519.5	56.0	Фізика: 171.0 Матем.: 180.0 Укр.м.: 150.0	—	-/-	+	—	—	-/+	6/0
95	Мельничук Андрій Васильович	516.5	50.0	Фізика: 169.5 Матем.: 170.0 Укр.м.: 169.0	—	-/-	+	—	—	-/+	5/0
96	Онищенко Денис Олегович	516.0	52.5	Фізика: 171.0 Матем.: 171.0 Укр.м.: 162.0	—	-/-	+	—	—	-/+	7/0

Рисунок 2.7 - Фрагмент рейтингового списку абітурієнтів, що мали прохідний бал, але не подали оригіналів документів.

2.4 Формування вимог до бази даних та модулю, що її формує

Отже, як вже відомо, ІС "Конкурс" містить всю необхідну інформацію про вступника, а саме: ПІБ абітурієнта, суму балів ЗНО, бал атестату про середню освіту, інформацію про пільги, а також наявність оригіналів документів. Окрім того, інформаційна система має інформацію щодо ВНЗ, факультету (інституту), спеціальності та ОКР, що вказані у заяві на вступ. А отже, з цими даними можна провести ряд експериментів, що дозволять визначити ВНЗ-конкуренти.

Для того щоб мати можливість сформувати статистичні дані, необхідно мати базу даних, до якої можна робити запити на вибірку та формувати статистику, а також механізм, що транслюватиме дані з ІС "Конкурс" до бази даних. Перш ніж почати розроблювати структуру бази даних і механізм завантаження даних, необхідно виявити вузькі місця, які можуть похитити дані та привести до помилкових висновків.

Раніше згадувалося, що у даному дослідженні конкуренція ВНЗ визначатиметься не тільки за кількістю заяв, до яких прикріплені оригінали документів, а й за числом осіб, які подають оригінали документів. Проблема полягає в тому, що дані з ІС «Конкурс» не можуть однозначно вказати на особу, тобто при виконанні запиту у базу з використанням імені абітурієнтів, ми отримаємо дані усіх осіб з цим ім'ям, що внесені в базу. Таким чином, існує необхідність вибору правильних критеріїв, даних та їх комбінацій, які б дозволили уникнути колізій під час дослідження.

Проте оскільки можливі випадки співпадіння цих параметрів у кількох осіб, імені абітурієнта чи комбінації імені та по батькові недостатньо для встановлення особи. Тобто необхідно розглянути інші види даних про абітурієнтів, що містяться у інформаційній системі, та визначити ті, які можна застосувати у даній роботі.

ІС «Конкурс» дає можливість використовувати суму балів сертифікатів ЗНО вступника, як критерій відбору особи. Згідно з [7], абітурієнт має право скласти не більше чотирьох тестувань з різних предметів і подати документи на різні спеціальності, але не більше трьох в одному ВНЗ, та не більше ніж у п'ять ВНЗ. Оскільки у різних навчальних закладів – різні вимоги щодо необхідних сертифікатів ЗНО, то навіть у одного абітурієнта цей параметр може суттєво відрізнятись. Тому використання цього критерію недоцільне.

У загальному доступі у базі ІС «Конкурс» зберігається інформація про додаткові бали за дипломи Всеукраїнських олімпіад та конкурсу МАН. Проте такі дані не можуть бути використані у якості критерію, адже не кожен абітурієнт має подібні бонуси чи не завжди може їх використати, бо бали враховуються лише якщо диплом стосується профільного предмету, що є необхідним для вступу до того чи іншого ВНЗ. Тобто для одного й того ж абітурієнта цей параметр може бути наявним чи відсутнім в залежності від ВНЗ, куди він подає документи.

Аналогічний висновок можна зробити про використання додаткових балів за диплом випускника факультету довузівської підготовки, оскільки цей бал є релевантним тільки для окремого ВНЗ, а отже може стати причиною недостовірного результату під час пошуку інформації про абітурієнта.

Використання такого параметру, як право на позаконкурсний та позачерговий вступ, що зберігається в базі, теж не є доцільним, оскільки цей критерій може приймати тільки два значення і існує вірогідність того, що у двох різних людей з абсолютно однаковими ПІБ ці параметри будуть теж однаковими.

Другим підходящим критерієм для даного дослідження є середній бал атестата про середню освіту абітурієнта, так як цей параметр може набувати будь-якого значення від 0.0 до 60.0 і вірогідність того, що такий параметр у двох різних людей з однаковими ПІБ співпаде, є невисокою. Також, на відміну від суми балів сертифікатів ЗНО, середній бал атестата не змінюється в залежності від обраного ВНЗ, факультету та спеціальності. Отже, доцільним є

використання середнього балу атестата про середню освіту в якості критерію пошуку та відбору абітурієнтів.

Під час проведення експериментів необхідно мати можливість не включати у вибірку дані, що не є цільовими. Як відомо, більшість вступників подає заяви на ОКР "Бакалавр" у ВНЗ III-IV рівня акредитації. А отже, база даних, у яку будуть завантажувати дані з ІС "Конкурс" повинна мати можливість виключити інформацію про вступників, що подали заяви на вступ до ВНЗ I-II рівня акредитації, та тих, хто бажав отримати ОКР "Магістр", "Спеціаліст", "Молодший спеціаліст" та всі інші ОКР, окрім "Бакалавр".

Виходячи з того, що деякі вступники мають право на позачерговий або позаконкурсний вступ до ВНЗ, то дані про таких вступників також необхідно виключити. Це зумовлено тим, що серед таких абітурієнтів є власний конкурс, а також, як правило, за таких вступників ВНЗ не конкурують.

2.5 Основні засади кластерного аналізу

Оскільки однією з цілей даної роботи є визначення тих ВНЗ, до яких найчастіше подають документи сильні абітурієнти, необхідно провести експеримент із використанням методів кластерного аналізу. Розглянемо основні теоретичні положення.

Нехай є вибірка об'єктів дослідження $s = \{o^{(1)}, \dots, o^{(N)}\}$, яка сформована в результаті відбору деяких представників генеральної сукупності. Потрібно сформуванати $K \geq 2$ класів (груп об'єктів); число класів може бути як вибрано заздалегідь, так і не задано (в останньому випадку оптимальну кількість кластерів має бути визначено автоматично). Кожен об'єкт генеральної сукупності описується за допомогою набору змінних X_1, \dots, X_n .

Набір $X = \{X_1, \dots, X_n\}$ може включати змінні різних типів (кількісні та якісні, під якими будемо розуміти номінальні і булеві, а також порядкові). Нехай D_j позначає безліч значень змінної X_j . Позначимо через $x = x(o) = x_1(o), \dots, x_n(o)$ набір спостережень змінних для об'єкта o , де $x_j(o)$ є

значення змінної X_j для даного об'єкта. Відповідний вибірці набір спостережень змінних будемо представляти у вигляді таблиці даних V з N рядками і n стовпцями: $V = \{ x_j^{(i)} \}$, $i=1,2,\dots,N$, $j=1,2,\dots,n$; при цьому значення $x_j^{(i)}$, що знаходиться на перетині i -го рядка і j -го стовпця відповідає спостереженню j -й змінної для i -го об'єкта: $x_j^{(i)} = X_j(o^{(i)})$. У деяких завданнях вихідна інформація являє собою таблицю попарних відстаней між об'єктами. Можна виділити такі основні етапи кластерного аналізу:

1. Формування системи змінних. Часто дослідник не може з упевненістю сказати, які саме змінні дійсно важливі для аналізу, тому прагне включити якомога більше потенційно інформативних факторів. Нерідко потрібно попередньо вибрати з вихідної множини змінних найбільш ефективну підсистему. Крім того, в деяких завданнях доцільно трансформувати вихідні змінні так, щоб утворити нові, більш інформативні показники («feature extraction»). Щоб уникнути «домінування» змінних з великим масштабом виміру, проводять попереднє нормування вихідних змінних.

2. Визначення способу обчислення відстані між об'єктами або групами об'єктів. Цей спосіб повинен відображати специфіку розв'язуваної прикладної задачі. Для кожної пари об'єктів $o^{(i)}$ і $o^{(l)}$ позначимо відстань між ними через $(o^{(i)}, o^{(l)})$, де i не дорівнює l . Наприклад, у випадку безперервних змінних може бути задане евклідову відстань:

$$(o^{(i)}, o^{(l)}) = \sqrt{\sum_{j=1}^n (x_j^{(i)} - x_j^{(l)})^2}. \quad (2.1)$$

Щоб виключити вплив сильних лінійних кореляцій між змінними, застосовують відстань Махалобіса:

$$(o^{(i)}, o^{(l)}) = (\mathbf{x}^{(i)} - \mathbf{x}^{(l)})^T \Sigma^{-1} (\mathbf{x}^{(i)} - \mathbf{x}^{(l)}) \quad (2.2)$$

де $\mathbf{x}^{(i)}$ і $\mathbf{x}^{(l)}$ – вектора-стовпці значень змінних для відповідних об'єктів, Σ – коваріаційна матриця (оцінена за вибіркою, або вважається відомою апріорі). Для номінальних змінних може використовуватися відстань

Хеммінга. Для груп об'єктів також визначається спосіб знаходження відстані, наприклад, за принципом «далекого сусіда», «ближнього сусіда» та ін. Принцип «далекого сусіда» виправданий у випадку, коли є апріорна інформація про те, що таксони мають компакту сферичну форму. Принцип «ближнього сусіда» є сенс застосовувати, якщо відомо, що таксони можуть мати «видовжену» форму або концентрично розташовані.

3. Угрупування об'єктів. На цьому кроці проводиться створення груп об'єктів. Розбиття на групи може бути «жорстким» (формується розбиття на безліч об'єктів), а може бути і «нечітким» (обчислюється ступінь приналежності кожного об'єкта до груп). Розглянемо угрупування першого типу: нехай сформовано згруповане рішення $G = \{C^{(1)}, C^{(2)}, \dots, C^{(K)}\}$, де $C^{(k)} \subseteq \{o^{(1)}, o^{(2)}, \dots, o^{(N_k)}\}$, N_k - число об'єктів, що входять в k -й кластер, $k=1, 2, \dots, K$. Під згрупованню вирішальною функцією будемо розуміти відображення $f: S \rightarrow \{1, 2, \dots, K\}$.

4. Представлення результатів. Потрібно отримати простий і інформативний опис отриманих кластерів. Часто для такого опису вибирається «типовий об'єкт» або визначається набір усереднених по групі показників. Використовується також опис у вигляді набору таксонів. Під таксоном будемо розуміти підобласть простору змінних мінімального обсягу, що має деяку задану форму і містить точки відповідної групи.

5. Визначення якості отриманого угрупування. Фахівцю прикладної області необхідно упевнитися в тому, що сформовані групи дійсно відображають внутрішні закономірності, характерні для розв'язуваної задачі, сприяють досягненню цілей аналізу, допомагають відкрити нові властивості досліджуваних об'єктів. Існують також більш формальні способи перевірки якості, пов'язані з перебуванням ймовірності випадкового утворення груп, яку можна обчислити в рамках тієї або іншої моделі розподілу (з перевіркою статистичних гіпотез про однорідність спостережень різних класів); з бутстреп-методом; з обчисленням різних показників якості (внутрішньогрупового розкиду, індексу Гудмана-Крускала; Ранда; З-індексу і т.д.)

На даний час існує декілька підходів до вирішення завдання кластерного аналізу, які засновані на різних уявленнях про завдання, використанні специфічною для кожної предметної області додаткової інформації і т.д. Коротко перерахуємо найбільш часто використовувані підходи.

- Імовірнісний підхід. Передбачається, що кожен об'єкт генеральної сукупності належить одному з K класів, проте номери класів безпосередньо не спостерігаються. Об'єкти обираються з генеральної сукупності випадково і незалежно, тому змінні, що описують об'єкти, випадкові. Для кожного класу визначено імовірнісний розподіл заданого сімейства; параметри розподілу невідомі. Наявна вибірка спостережень являє собою реалізацію суміші розподілів. Необхідно визначити найбільш правдоподібні значення параметрів, відновивши закон розподілу для кожного класу. Існують наближені алгоритми розщеплення суміші (EM-алгоритм). Перевірка значущості поділу може бути проведена з використанням апарату перевірки статистичних гіпотез. Відомі також алгоритми, засновані на непараметричних оцінках щільності.

- Підхід, який використовує аналогію з центром ваги. Для кожної групи визначається вектор середніх значень показників, що інтерпретується як «центр тяжкості» групи. Використовується критерій внутрішньогрупового розсіювання:

$$d(G) = \sum_{k=1}^K \sum_{l=1}^{N_k} \sum_{j=1}^n (x_{ij}^{(kl)} - \bar{x}_{kl})^2. \quad (2.3)$$

Оптимальна угруповання, при заданому K , відповідає мінімальному значенню критерія. В алгоритмі K -середніх згрупованих рішень формується динамічно з деякою вихідної угруповання шляхом поетапного перерозподілу об'єктів в групи з найближчими центрами тяжіння. Це перерозподіл йде до отримання стійкого поділу. Аналогічна методика використовується також в алгоритмі FOREL.

- Підхід, заснований на теорії графів. Найбільш відомий алгоритм цього сімейства – алгоритм найкоротшого незамкнутого шляху. Попередньо

будується мінімальне кістяк графа, в якому вершини відповідають об'єктам, а ребра мають довжину, рівну відстані між відповідними об'єктами. Для утворення кластерів з побудованого дерева видаляються ребра максимальної довжини.

- Ієрархічний підхід. Даний напрямок також має відношення до теоретико-графових підходів. Результати угруповання представляються у вигляді дерева угруповання (дендрограмми). Алгоритми, засновані на цьому підході, можна розділити на агломеративні (поетапно об'єднують найближчі групи або об'єкти) і дівізимні (в яких поетапно здійснюється поділ вихідної групи на найбільш віддалені підгрупи; ті в свою чергу також поділяються на підгрупи і т.д.). Згрупороване рішення – це вкладена ієрархія підгруп.

- Підхід, заснований на понятті найближчого сусіда. Угруповання здійснюється послідовно шляхом приписування об'єкту кластеру, в якому знаходиться найближчий об'єкт, за умови, що відстань до об'єкта не перевищує заданий поріг. Існують різні варіанти визначення відстані; при визначенні міри близькості може враховуватися і розташування інших сусідніх точок.

- Нечіткі алгоритми кластерного аналізу. При використанні даного підходу передбачається, що кожен кластер являє собою нечітку множину об'єктів. До найбільш популярних алгоритмам цього сімейства можна віднести алгоритм нечітких С-середніх.

- Підхід, що використовує штучні нейронні мережі, заснований на аналогії з процесами, що відбуваються в біологічних нейронних системах. Відомо велике число алгоритмів даного сімейства. Типова архітектура являє собою одношарову мережу, в якій кожен нейрон відповідає деякому кластеру. У процесі навчання мережі відбувається ітеративна зміна передавальних ваг між вхідними та вихідними вузлами мережі; тим самим здійснюється пошук оптимального значення критерію угруповання. Нейронні мережі дозволяють ефективно використовувати паралельні методи обчислень. Привабливим властивістю самоорганізованих мереж Кохонена є те, що вони формують наочне двовимірне відображення безлічі об'єктів. Існує певна подібність у

процесі угруповання між алгоритмами, заснованими на нейронних мережах та деякими класичними методами кластерного аналізу.

Для проведення експерименту у роботі буде використаний підхід, заснований на понятті найближчого сусіда. Даний метод було обрано, оскільки апріорі немає можливості дізнатись яку кількість кластерів буде отримано у результаті. А використовуючи цей метод, можна буде чітко визначити кластери ВНЗ.

2.5 Висновки до другого розділу

У даному розділі визначено поняття ВНЗ-конкурентів та сформовано вектори розгляду та вимоги до експериментів, за результатами яких можна визначити такі навчальні заклади. Також у розділі зроблено аналіз даних, що доступні на ІС "Конкурс". Також у цій частині роботи було сформовано вимоги до БД і критерії визначення унікальних осіб. Окреслено основні засади та методи кластерного аналізу. Для побудови кластеру у роботі буде використано підхід найближчого сусіда, у якому відстань буде визначено за формулою евклідової відстані.

3 РОЗРОБКА МОДУЛЯ ЗАВАНТАЖЕННЯ ТА ЗАПИСУ ДАНИХ З ІС "КОНКУРС" ДО БД

3.1 Формування бази даних рейтингів вступників

Однією з ключових задач роботи є формування структури бази даних, до якої необхідно завантажити дані із ІС "Конкурс". Для цієї задачі доцільно використати реляційну базу даних під управлінням системи управління базами даних MSSQL.

Дана система управління надає широкі можливості для швидкого створення структури бази даних (за допомогою інтерфейсу користувача), а також різноманітні механізми для аналізу скриптів, що формують необхідну статистику. Для того щоб найбільш ефективно сформувати необхідну для аналізу статистику, необхідно створити структуру бази даних, яка б моделювала предметну область. Структуру розробленої БД зображено на рис. 3.1.

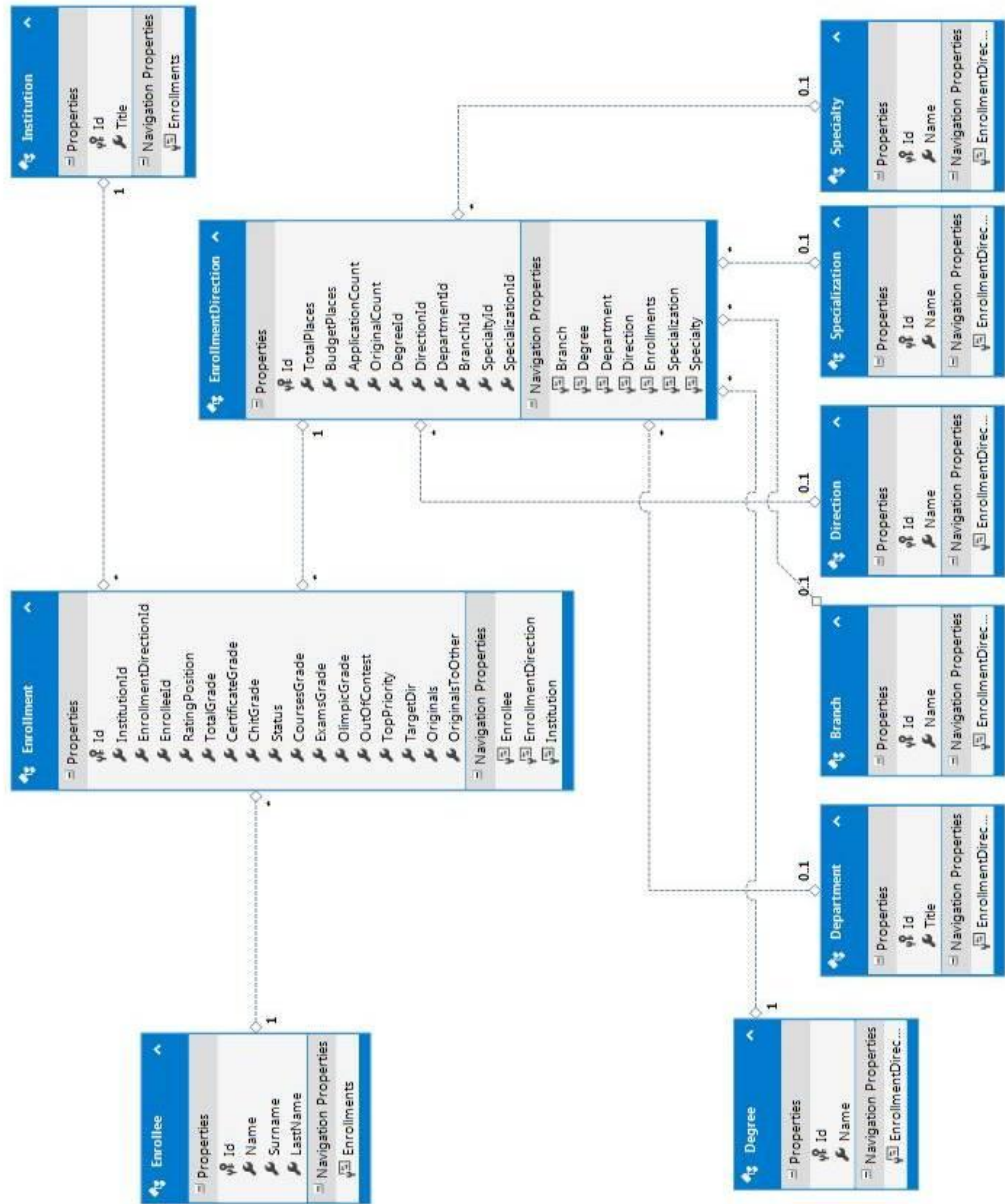


Рисунок 3.1 - Структура створенної БД

Отже, створена структура складається із 10 таблиць різного розміру й структури. У таблиці [Enrollement] зберігається інформація про заяви вступників. Таблиця складається з таких колонок: Id - (int, not null) унікальний ідентифікатор заяви. Колонка є первинним ключем.

- InstitutionId - (int, not null) зовнішній ключ на таблицю [Institution]. У колонці зберігається інформація про ВНЗ, у який вступник подав заяву.
- EnrollmentDirectionId - (int, not null) зовнішній ключ на таблицю розв'язки [EnrollmentDirection].
- EnrolleeId - (int, not null) зовнішній ключ на таблицю [Enrollee]. У колонці зберігається інформація про вступника, що подав заяву.
- RatingPosition - (smallint, not null) позиція заяви у рейтинговому списку.
- TotalGrade - (real, not null) загальний бал.
- CertificateGrade - (real, not null) сумарний бал сертифікатів ЗНО.
- ChitGrade - (real, not null) бал атестату про середню освіту.
- CoursesGrade - (real, not null) додатковий бал за курси доуніверситетської підготовки.
- ExamsGrade - (real, not null) бали за додаткові вступні іспити.
- OlympicGrade - (real, not null) додаткові бали за призові місця олімпіад.
- OutOfContest - (bit, not null) параметр вказує на право поза конкурсного вступу.
- Originals - (bit, not null) параметр вказує на наявність оригіналів документів.
- OriginalsToOther - (bit, not null) параметр вказує на наявність оригіналів на інші напрямки/ВНЗ.

У таблиці [Enrollee] зберігається інформація про вступників. Таблиця складається з таких колонок:

- Id - (int, not null) унікальний ідентифікатор вступника. Колонка є первинним ключем.
- Name - (nvarchar(30), not null) ім'я вступника.
- Surname - (nvarchar(30), not null) прізвище вступника.

- Lastname - (nvarchar(30), not null) по-батькові вступника.

У таблиці [Institution] зберігається інформація про ВНЗ. Таблиця складається з таких колонок:

- Id - (int, not null) унікальний ідентифікатор ВНЗ. Колонка є первинним ключем.
- Title - (nvarchar(240), not null) назва ВНЗ.

У таблиці [EnrollmentDirections] зберігається детальна інформація про спеціальність, напрям, ОКР, обсяг конкурсу та ін. Таблиця складається із таких колонок:

- Id - (int, not null) унікальний ідентифікатор. Колонка є первинним ключем.
- DegreeId - (int, not null) зовнішній ключ на таблицю [Degree]. У колонці зберігається інформація про ОКР, що вказаний у заяві на вступ.
- DirectionId - (int, not null) зовнішній ключ на таблицю [Direction]. У колонці зберігається інформація про напрям, що вказаний у заяві на вступ.
- DepartmentId - (int, not null) зовнішній ключ на таблицю [Department]. У колонці зберігається інформація про факультет, що вказаний у заяві вступника.
- BranchId - (int, not null) зовнішній ключ на таблицю [Branch]. У колонці зберігається інформація про галузь, що вказана у заяві вступника.
- SpecialtyId - (int, not null) зовнішній ключ на таблицю [Specialty]. У колонці зберігається інформація про спеціальність, що вказана у заяві вступника.
- SpecializationId - (int, not null) зовнішній ключ на таблицю [Specialization]. У колонці зберігається інформація про спеціалізацію, що вказана у заяві вступника.
- TotalPlaces - (smallint, not null) загальна кількість місць.
- BudgetPlaces - (smallint, not null) кількість бюджетних місць.
- ApplicationCount - (smallint, not null) кількість поданих заяв.

- OriginalCount - (smallint, not null) кількість поданих оригіналів документів.

У таблиці [Degree] зберігається інформація про ОКР та складається із наступних колонок:

- Id - (int, not null) унікальний ідентифікатор ОКР. Колонка є первинним ключем.
- Name - (nvarchar(60), not null) назва ОКР.

У таблиці [Branch] зберігається інформація про галузі та складається із наступних колонок:

- Id - (int, not null) унікальний ідентифікатор галузі. Колонка є первинним ключем.
- Name - (nvarchar(40), not null) назва галузі.

У таблиці [Department] зберігається інформація про факультети ВНЗ та складається із наступних колонок:

- Id - (int, not null) унікальний ідентифікатор факультету. Колонка є первинним ключем.
- Title - (nvarchar(240), not null) назва факультету.

У таблиці [Direction] зберігається інформація про напрями та складається із наступних колонок:

- Id - (int, not null) унікальний ідентифікатор напрямку. Колонка є первинним ключем.
- Name - (nvarchar(40), not null) назва напрямку.

У таблиці [Specialization] зберігається інформація про спеціалізації та складається із наступних колонок:

- Id - (int, not null) унікальний ідентифікатор спеціалізації. Колонка є первинним ключем.
- Name - (nvarchar(40), not null) назва спеціалізації.

У таблиці [Specialities] зберігається інформація про спеціальності та складається із наступних колонок:

- Id - (int, not null) унікальний ідентифікатор спеціальності. Колонка є первинним ключем.
- Name - (nvarchar(60), not null) назва спеціальності.

Структуру БД буде сформовано за допомогою спеціальних процедур. Лістинги процедур наведено у "Додатку Б".

3.2 Завантаження даних абітурієнтів у базу

У попередньому підрозділі була описана структура бази даних, що необхідна для формування статистики та подальшого її аналізу. Наразі постає завдання завантаження даних із відкритого джерела ІС "Конкурс".

У якості базового механізму завантаження необхідно взяти модуль, що був розроблений у бакалаврській роботі для пошуку інформації про абітурієнта на ІС "Конкурс".

Структурна схема розробленого модуля зображено на рис. 3.1.

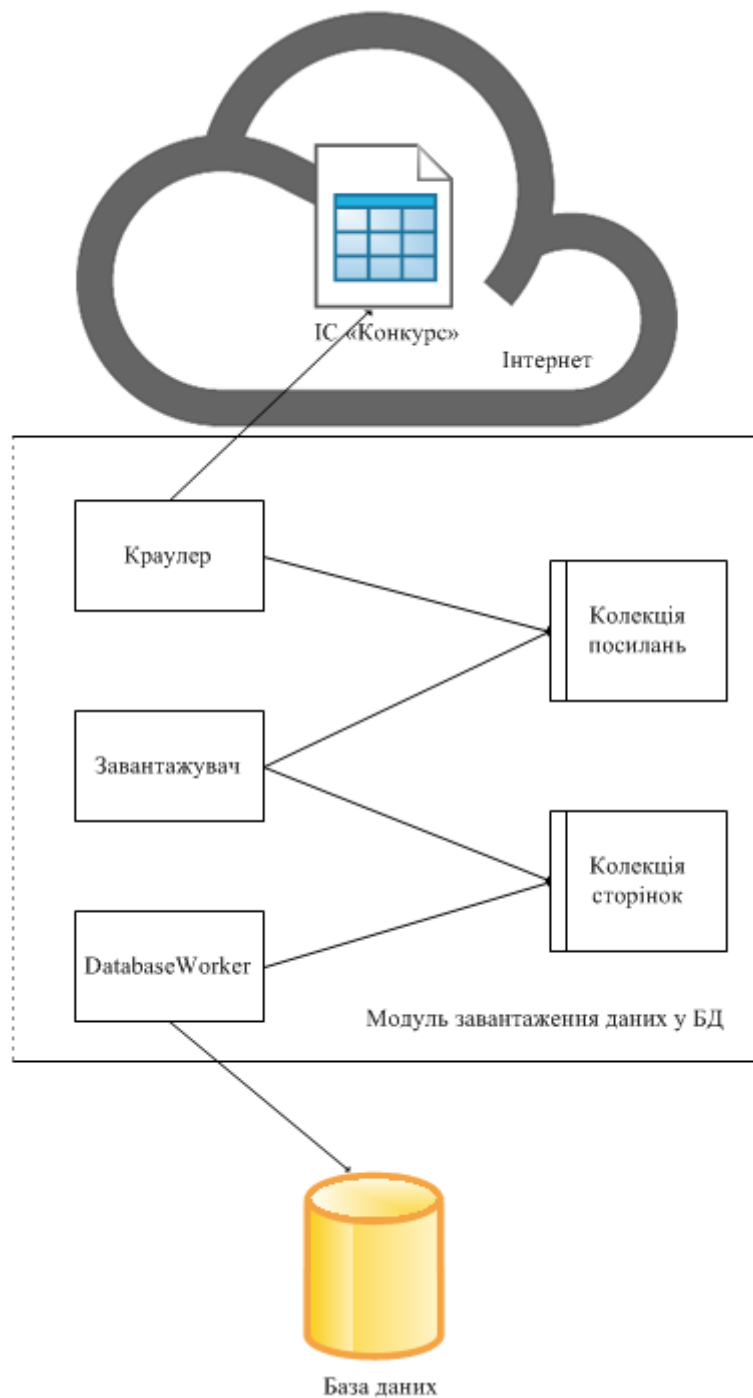


Рисунок 3.1 - Структурна схема розробленого модуля

Завантаження даних складається з трьох етапів, організованих у конвеєр (FetchPipeline):

1. Кравлинг сайту з метою пошуку сторінок для завантаження (сторінки з даними конкурсу).
2. Завантаження сторінок з даними конкурсу.

3. Обробка сторінок з даними конкурсу з метою вилучення та додання у базу даних.

Під час дії конвеєру запускаються його окремі частини, які будуть працювати одночасно – кожен етап конвеєру використовує один або більше потоків. Для передачі даних між частинами конвеєру використовуються спеціальні контейнери у складі .Net Framework з простору імен System.Collections.Concurrent такі як BlockingCollection.

Загальний алгоритм роботи модуля зображено у вигляді UML-діаграми послідовностей (sequence diagram). На рис. 3.2 – графічне зображення діаграми:

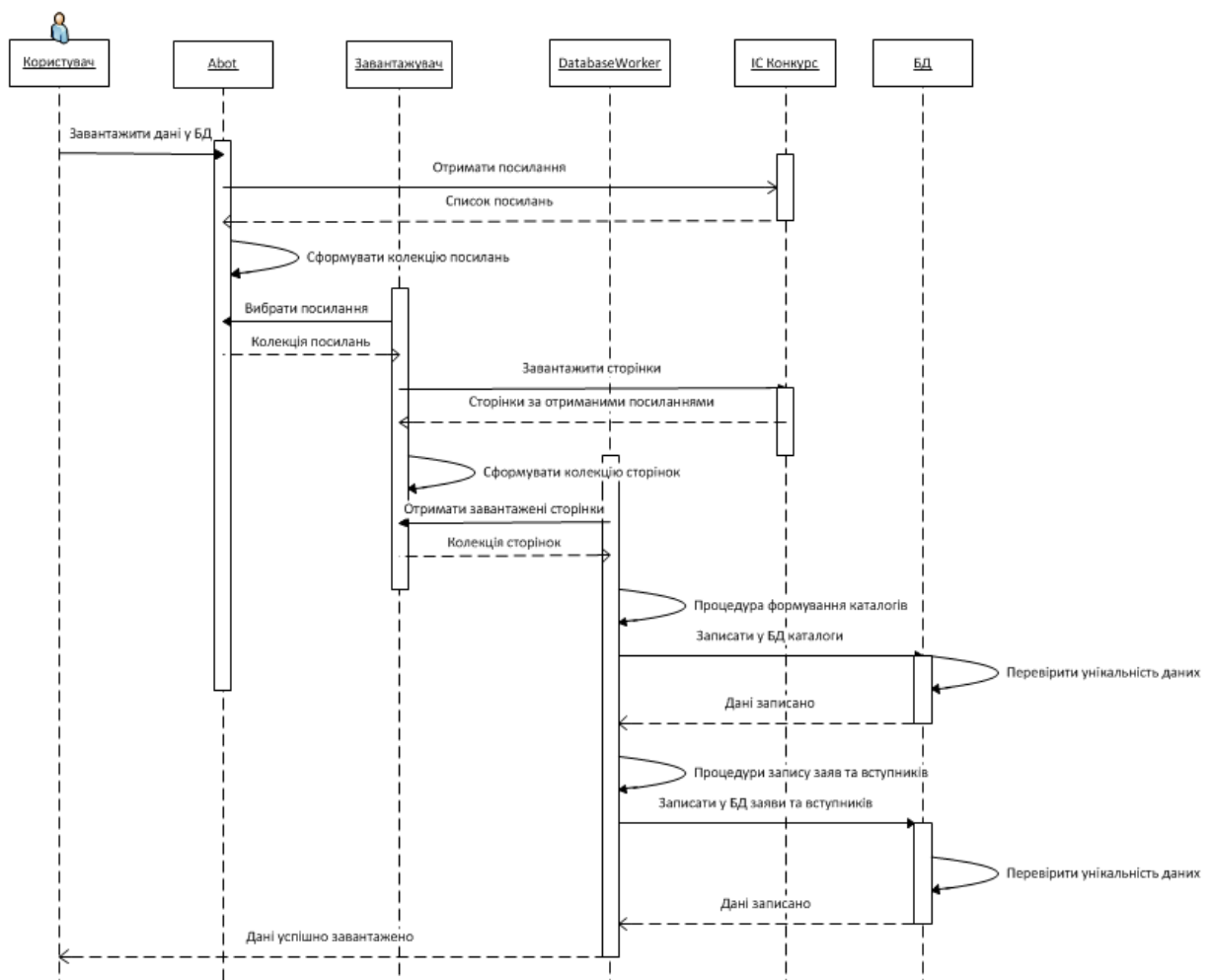


Рисунок 3.2 - UML-діаграма роботи модуля завантаження даних у БД

3.2.1 Краулер (CrawlerWrapper/Abot)

Робота підсистеми, що розроблялась, починається з проходження та завантаження потрібної інформації з веб-сайту <http://vstup.info>. Для того щоб вилучити список абітурієнтів, необхідно пройти п'ять рівнів посилань. Під час проходження кожного рівня необхідно запобігти переходу за посиланнями та не включати обробку непотрібної інформації.

Процес проходження по веб-сайту реалізується за допомогою краулера – набору скриптів (алгоритмів), за допомогою яких відбувається перехід по сторінкам веб-ресурсу та збір необхідної інформації. У даній роботі використовувався краулер Abot.

Abot – це краулер із відкритим вихідним кодом C#, створений для швидкої і гнучкої роботи з веб-ресурсами. Він використовує набір інструментів (багатопоточність, HTTP-запити і т.д.) задля проходження по контенту сайту та отримання потрібної інформації.

Abot має досить широкий функціонал та дозволяє, крім проходження по сторінках ІС "Конкурс", ще й завантажувати інформацію звідти. Але використання цього краулера у якості завантажувача цільових сторінок для даної роботи буде неефективним, адже деревовидна структура ІС "Конкурс" розширюється дуже швидко. До того ж, для реалізації модуля застосовується багатопоточність. Тому доцільно буде використовувати Abot лише для переходу за посиланнями, колекціонування посилань з областей веб-ресурсу, що розглядається, та відсіювання областей, які не цікаві для задачі, що вирішується.

Логіка роботи з краулером загорнута у спеціальний клас (CrawlerWrapper), що приховує роботу з краулером та відкриває більш простий інтерфейс керування. Для направлення краулера на цільову сторінку використовуються білі та чорні списки посилань, по яким краулеру відповідно дозволено або заборонено переходити. Ці списки складаються з регулярних виразів, що описують кожне з посилань.

3.2.2 Завантажувач (ConcurrentWebPageDownloader / WebClient)

Так як для завантаження великої кількості цільових сторінок потрібно використовувати окремий механізм, такий механізм було створено. У складі .Net Framework у просторі імен System.Net є клас, призначений для завантаження веб сторінок – WebClient. Один WebClient може одночасно завантажувати одну сторінку. Для досягнення високих швидкостей завантаження сторінок використовується декілька WebClient, кожному з яких видається нове завдання на завантаження після виконання попереднього.

Завантажувач (ConcurrentWebPageDownloader) працює багатопоточно. Для завантаження веб-сторінок використовуються дві колекції BlockingCollection<T> – одна містить вільні веб клієнти (WebClient), а інша – адреси сторінок, які треба завантажити. Веб-клієнти (WebClient) асинхронно завантажують веб-сторінки, адреси яких містяться у іншій колекції. У той самий час, краулер Abot заповнює колекцію адрес веб-сторінок. У даній реалізації було використано колекцію BlockingCollection<T>, оскільки саме цей інструмент підтримує багатопоточність (є потокозахисним) та є найбільш зручним інструментом для динамічного керування його вмістом. Якщо у колекції веб-клієнтів є вільний клієнт, то він вилучається з колекції для завантаження веб-сторінки. Коли завантаження закінчиться, веб-клієнт згенерує подію, під час якої він повернеться до колекції вільних веб-клієнтів та згенерує подію завантажувача, яка сповістить про те, що сторінку було завантажено.

Оскільки текстова інформація представлена українською мовою в кодуванні Кирилиця (сторінка кодів 1251) , то необхідно налаштувати краулер на роботу з цим кодуванням. Для цього обробка його запитів налаштовується таким чином, що б він декодував сторінки з кодування 1251 у кодування Unicode, яке використовується середою .NET Framework для зберігання та відображення тексту. Це дозволить запобігти отриманню, обробці та відображенню інформації у некоректному вигляді.

3.2.3 Додання у БД (DatabaseWorker + БД)

Для заповнення бази даних використовуються два модулі: один у програмі (DatabaseWorker), а інший – у БД.

Клас DatabaseWorker є останньою, третьою ланкою конвеєра. На вході він отримує вміст та метадані цільових сторінок конкурсу. За допомогою регулярних виразів виокремлюємо корисні дані та додаємо їх до БД за допомогою процедур, що зберігаються. Загальні дані про напрям та інформацію про цю сторінку додає у БД процедура InsertCatalogs. Іншим регулярним виразом виокремлюються усі подані заяви на цій сторінці та додаються в БД за допомогою процедури InsertEnrollment.

Дані процедури, що зберігаються, реалізують увесь функціонал додання у БД, таким чином є її інтерфейсом на запис. Це дозволяє виконати принцип інкапсуляції для бази даних. Також, їх використання дозволяє зменшити навантаження на канал програма-БД, що значно підвищує швидкість роботи.

Процедури, що зберігаються, реалізовані на мові Transact SQL з використанням оператора MERGE, який, у свою чергу, дозволяє зменшити накладні витрати.

3.2.4. Конвеєр

Конвеєр FetchPipeline об'єднує три частини модуля завантаження даних. Він відповідає за запуск (створення та ініціалізацію екземплярів) кожного з етапів та передачу даних між ними. Конвеєр містить події, через які сповіщує про повне завершення роботи кожного з етапів. Конвеєр також має властивості, через які можливо відстежувати прогрес будь-якого з етапів або його частин.

Конвеєр починає працювати з викликом функції GoAsync. Під час запуску перевіряється поточний стан конвеєру, можливе його часткове приведення в роботу. Кожний етап конвеєру працює асинхронно, у одному або декількох потоках. Кожен з етапів доповідає за виконання свого завдання через події, на які підписаний конвеєр, що дозволяє йому реагувати на них та передавати дані далі.

Через швидке розширення дерева сайту нам потрібно знизити навантаження на краулер та делегувати роботу по завантаженню сторінок спеціалізованому рішенню – завантажувачу. Ми зупиняємо роботу краулера на сторінках з посиланнями «конкурс», та вручну вилучаємо посилання з тексту сторінки.

Як видно з рисунків вище, сторінки з рейтингами абітурієнтів не мають будь-якої інформації про спеціальність. А отже, додавач не зможе визначити спеціальність, з яким пов'язаний завантажений рейтинг. Але ці дані є на попередній сторінці, на якій наведена таблиця з інформацією про факультети, що є в обраному ВНЗ, напрямки та спеціальності, які ці факультети мають. Для кожного такого рядка є посилання "конкурс", після відкриття якого з'являється рейтинг абітурієнтів, що подали документи на вступ до конкретної спеціальності, конкретного факультету. Отже, необхідно запам'ятовувати, рейтинг якої спеціальності завантажується. Це можна реалізувати за допомогою посилання "конкурс", яке прив'язане до факультету та ВНЗ з одного боку, та з іншого, до рейтингу абітурієнтів. Також нам важливо запам'ятати кількість заяв що було подано на цей напрям (посилання конкурс).

Для виконання цих дій необхідний проміжний етап, що реалізований у самому конвеєрі – при отриманні цільової сторінки с посиланнями «конкурс» та передачі на завантаження завантажувачеві. Ці дані прикріплюються до інформації про цільову сторінку та передаються далі.

3.3 Висновки до третього розділу

У даному розділі описана структура БД, у яку будуть записуватись дані із ІС "Конкурс". Загалом БД складається з 10 таблиць та має розмір 276 МБ. Також у розділі описано реалізацію модуля, що виконує завантаження даних та їх запис до БД, наведені використані програмні рішення. Розроблений модуль також було оптимізовано для найбільш ефективного завантаження.

4 РЕЗУЛЬТАТИ ПРОВЕДЕНИХ ЕКСПЕРИМЕНТІВ

4.1 Тестування роботи модуля

4.1.1 Перевірка правильності завантажених даних у БД

Розроблену підсистему необхідно ретельно перевірити та протестувати, щоб упевнитись, що на виході ми отримуємо правильний результат. Також необхідно упевнитись, що застосовані засоби програмування максимально збільшують продуктивність роботи підсистеми. Головне вікно системи зображено на рис. 4.1:

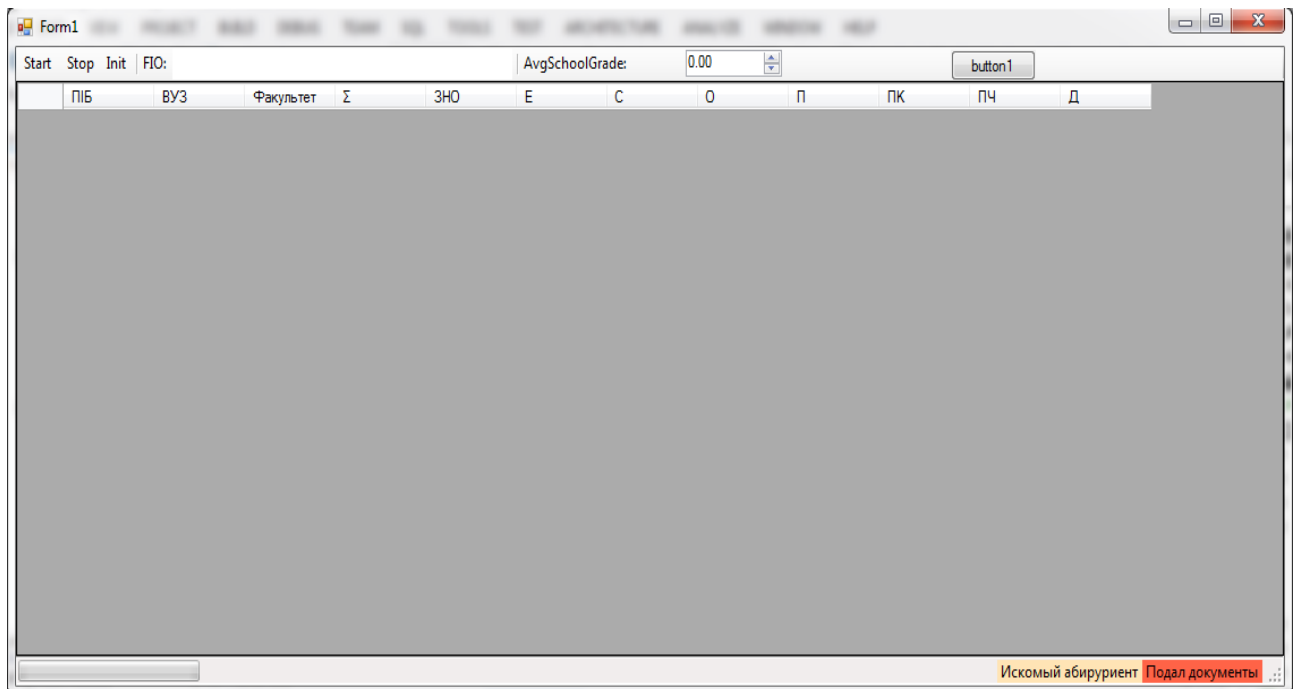


Рисунок. 4.1 - Головне вікно модуля

По-перше, необхідно упевнитись, що за допомогою обраних критеріїв відбору унікальних вступників формується БД, що має коректні дані, які в подальшому можна використовувати для проведення експериментів. Для початку перевіримо результат пошуку лише за прізвищем.

Використаємо комбінацію ПІБ "Мельник Андрій Володимирович" для пошуку. Результат наведено на рисунку 4.2:

ПІБ	ВУЗ	Факультет	Σ	ЗНО	Е	С	О	П	ПК	ПЧ	Д
Мельник Андрій Володимирович	Львівський національний аграрний університет	Механіки та енергетики	124	0.00	124.00	0.00	-	-	+	-	+
Мельник Андрій Володимирович	Луцький інститут розвитку людини Відкритого міжнародного Університету розвитку людини "Україна"		573.2	415.00	0.00	158.20	-	-	-	-	+
Мельник Андрій Володимирович	Львівський національний університет ветеринарної медицини та біотехнологій імені С.З. Гжицького	факультет ветеринарної медицини	351.9	0.00	170.00	181.90	-	-	-	-	+
Мельник Андрій Володимирович	Національний університет харчових технологій	інженерної механіки та пакувальної техніки	714	523.50	0.00	190.50	-	-	-	+	-
Мельник Андрій Володимирович	Національний авіаційний університет	Механіко-енергетичний факультет	714	523.50	0.00	190.50	-	-	-	+	-
Мельник Андрій Володимирович	Національний університет харчових технологій	енергетики і енергоменеджменту	714	523.50	0.00	190.50	-	-	-	+	-
Мельник Андрій Володимирович	Національний технічний університет України «Київський політехнічний інститут»	Теплоенергетичний факультет	733.1	523.50	0.00	190.50	-	19.10	-	+	+
Мельник Андрій Володимирович	Національний технічний університет України «Київський політехнічний інститут»	Міжуніверситетський медико-інженерний факультет	733	523.50	0.00	190.50	-	19.00	-	+	-
Мельник Андрій Володимирович	Національний авіаційний університет	Інститут аерокосмічних систем управління	714	523.50	0.00	190.50	-	-	-	+	-
Мельник Андрій Володимирович	Національний університет харчових технологій	енергетики і енергоменеджменту	714	523.50	0.00	190.50	-	-	-	+	-
Мельник Андрій Володимирович	Хмельницький університет управління та права	Юридичний факультет	757.4	565.00	0.00	192.40	-	-	-	-	+
Мельник Андрій Володимирович	Львівський національний університет імені Івана Франка	Юридичний факультет	757.4	565.00	0.00	192.40	-	-	-	-	-
Мельник Андрій Володимирович	Львівський національний університет імені Івана Франка	Географічний факультет	776.4	584.00	0.00	192.40	-	-	-	-	-
Мельник Андрій Володимирович	Київський національний університет імені Тараса Шевченка	Економічний факультет	733.8	540.50	0.00	193.30	-	-	-	+	-
Мельник Андрій Володимирович	Київський національний торговельно-економічний університет	Факультет обліку, аудиту та економічної кібернетики	733.8	540.50	0.00	193.30	-	-	-	+	-
Мельник Андрій Володимирович	Національний університет біоресурсів і природокористування України	Аграрного менеджменту	733.8	540.50	0.00	193.30	-	-	-	+	-
Мельник Андрій Володимирович	Національний університет біоресурсів і природокористування України	Економічний	733.8	540.50	0.00	193.30	-	-	-	+	-
Мельник Андрій Володимирович	Київський національний університет імені Тараса Шевченка	Економічний факультет	733.8	540.50	0.00	193.30	-	-	-	+	-
Мельник Андрій Володимирович	Національний університет біоресурсів і природокористування України	Технології виробництва і переробки продукції тваринництва	712.3	519.00	0.00	193.30	-	-	-	+	-
Мельник Андрій Володимирович	Київський національний університет будівництва і архітектури	Геоінформаційні системи і управління територіями	733.8	540.50	0.00	193.30	-	-	-	+	-
Мельник Андрій Володимирович	Київський національний університет будівництва і архітектури	Будівельний	733.8	540.50	0.00	193.30	-	-	-	+	+
Мельник Андрій Володимирович	Київський національний торговельно-економічний університет	Факультет фінансів та банківської справи	733.8	540.50	0.00	193.30	-	-	-	+	-
Мельник Андрій Володимирович	Київський національний торговельно-економічний університет	Факультет економіки, менеджменту і права	733.8	540.50	0.00	193.30	-	-	-	+	-
Мельник Андрій Володимирович	Київський національний університет будівництва і архітектури	Геоінформаційні системи і управління територіями	733.8	540.50	0.00	193.30	-	-	-	+	-

Рисунок 4.2 - Результат пошуку особи "Мельник Андрій Володимирович"

Бачимо, що було знайдено щонайменше шість осіб. Для того щоб запобігти кореляцій, як зазначалося раніше, ми використовуємо середній бал атестата про середню освіту. Отже, припустімо, нам необхідно знайти додаткову інформацію про людину із ПІБ Мельник Андрій Володимирович та середнім балом атестата про середню освіту – 54.50. Результат зображено на рисунку 4.3:

Мельник Андрій Володимирович	Національний університет харчових технологій	інженерної механіки та пакувальної техн...	714	523.50	0.00	54.50	-	-	-	+	-
Мельник Андрій Володимирович	Національний авіаційний університет	Механіко-енергетичний факультет	714	523.50	0.00	54.50	-	-	-	+	-
Мельник Андрій Володимирович	Національний університет харчових технологій	енергетики і енергоменеджменту	714	523.50	0.00	54.50	-	-	-	+	-
Мельник Андрій Володимирович	Національний технічний університет України «Київськ...	Теплоенергетичний факультет	733.1	523.50	0.00	54.50	-	19.10	-	+	+
Мельник Андрій Володимирович	Національний технічний університет України «Київськ...	Міжуніверситетський медико-інженерний...	733	523.50	0.00	54.50	-	19.00	-	+	-
Мельник Андрій Володимирович	Національний авіаційний університет	Інститут аерокосмічних систем управління	714	523.50	0.00	54.50	-	-	-	+	-
Мельник Андрій Володимирович	Національний університет харчових технологій	енергетики і енергоменеджменту	714	523.50	0.00	54.50	-	-	-	+	-

Рисунок 4.3 - Результат пошуку додаткової інформації про абітурієнта Мельника Андрія Володимировича із середнім балом атестата 54.50

Як видно із зображень вище, дані критерії можуть бути застосовані для відбору унікальних вступників, навіть якщо серед них зустрічаються люди із однаковими ПІБ.

По-друге, необхідно перевірити коректність даних, записаних у базу, тобто упевнитись у тому, що у БД потрапили всі необхідні дані без зайвої інформації.

Для того щоб виконати таку перевірку, можна виконати запит у створеній БД та порівняти із аналітичними даними, що відображаються у ІС "Конкурс".

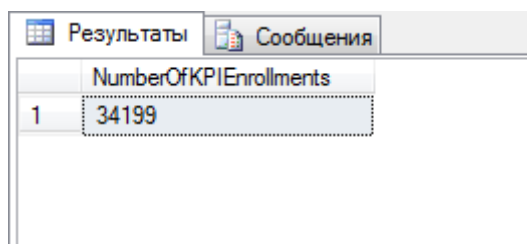
1) Перевіримо кількість поданих заяв на ОКР "Бакалавр" до НТУУ "КПІ" у 2014 році. Для перевірки у БД виконаємо запит:

```
USE AdmissionDB;
```

```
SELECT COUNT(DISTINCT et.Id) AS NumberOfKPIEnrollments  
FROM Enrollments AS et  
INNER JOIN Institutions AS i ON i.Id = et.InstitutionId  
INNER JOIN EnrollmentDirections AS ed ON ed.Id = et.EnrollmentDirectionId  
INNER JOIN Degrees ON dg ON dg.Id = ed.DegreeId
```

```
WHERE i.Title LIKE N'%Київський політехнічний інститут%' AND DegreeId = 3
```

Результат відображено на рис. 4.4:



	NumberOfKPIEnrollments
1	34199

Рисунок 4.4 - Кількість поданих заяв до НТУУ "КПІ", визначена у створеній БД

Дані із ІС "Конкурс" наведено у рис. 4.5:

Про ВНЗ	Статистика	Денна форма		Заочна форма		
		Всього	Мол. спец.	Бакалавр	Спеціаліст	Магістр
Подано заяв:		53072	0	44942	4459	3670
Допущено до конкурсу:		39921	0	34199	2584	3137
Серед допущених: заяв із правом позаконкурсного вступу:		588	0	552	15	21
Серед допущених: подано оригіналів документів:		9598	0	5705	2069	1823
Серед допущених: рекомендовано до зарахування:		137	0	10	92	35
Серед допущених: зараховано:		9028	0	5294	1998	1735
Ліцензований обсяг:		21854	0	10827	7053	3972
Бюджетних місць:		7473	0	4361	1682	1430
Загальний конкурс, осіб на місце:		1.83	0	3.16	0.37	0.79
Конкурс на бюджетні місця, осіб на місце:		5.34	0	7.84	1.54	2.19

Рисунок 4.5 - Дані ІС "Конкурс" про кількість допущених вступників до конкурсу НТУУ "КПІ" станом на 2014 рік

2) Визначимо загальну кількість бюджетних місць у НТУУ "КПІ".

Використаємо для перевірки наступний запит:

```
SELECT SUM(BudgetPlaces) as KPIBudgetPlaces
```

```
FROM EnrollmentDirections as ed
```

```
WHERE Id in (SELECT distinct ed.Id
```

```
FROM Enrollments AS et
```

```
INNER JOIN Institutions AS i ON i.Id = et.InstitutionId
```

```
INNER JOIN EnrollmentDirections AS ed ON ed.Id = et.EnrollmentDirectionId
```

```
INNER JOIN Degrees AS dg ON dg.Id = ed.DegreeId
```

```
WHERE i.Title LIKE N'%Київський політехнічний інститут%' AND DegreeId = 3
```

Результат відображено на рис. 4.6:

KPIBudgetPlaces	
1	4361

Рисунок 4.6 - Кількість бюджетних місць у НТУУ "КПІ" станом на 2014 рік

Дані із ІС "Конкурс" наведено у рис. 4.7:

Вступна кампанія 2014 » Національний технічний університет України «Київський політехнічний інститут» » Статистика

Сторінку оновлено 27/09/14 о 22:40 за даними ВНЗ з допомогою Єдиної державної електронної бази з питань освіти.

Про ВНЗ	Статистика	Денна форма	Заочна форма			
		Всього	Мол. спец.	Бакалавр	Спеціаліст	Магістр
Подано заяв:		53072	0	44942	4459	3670
Допущено до конкурсу:		39921	0	34199	2584	3137
Серед допущених: заяв із правом позаконкурсного вступу:		588	0	552	15	21
Серед допущених: подано оригіналів документів:		9598	0	5705	2069	1823
Серед допущених: рекомендовано до зарахування:		137	0	10	92	35
Серед допущених: зараховано:		9028	0	5294	1998	1735
Ліцензований обсяг:		21854	0	10827	7053	3972
Бюджетних місць:		7473	0	4361	1682	1430
Загальний конкурс, осіб на місце:		1.83	0	3.16	0.37	0.79
Конкурс на бюджетні місця, осіб на місце:		5.34	0	7.84	1.54	2.19

Рисунок 4.7 - Дані ІС "Конкурс" про обсяг бюджетних місць ОКР "Бакалавр" у НТУУ "КПІ" станом на 2014 рік

Отже, із наведених вище перевірок, видно, що створений завантажувач даних із ІС "Конкурс" виконав свою роботу вірно.

4.1.2 Оцінка продуктивності реалізованого модуля

У попередньому розділі було охарактеризовано програмні засоби, що використовувались для оптимізації роботи модуля завантаження даних у БД.

Під час конкурсного відбору абітурієнтів керівники відбіркових комісій можуть застосовувати розроблений модуль для оперативного визначення переміщень абітурієнтів між ВНЗ та факультетами. Враховуючи це, необхідно упевнитись, що розроблений модуль буде швидко виконувати усі операції по завантаженню даних. Якщо, наприклад, система буде завантажувати дані більше ніж добу, то співробітники відбіркових комісій не зможуть повноцінно використовувати розроблений модуль.

Для того щоб зробити висновки щодо продуктивності роботи модуля, необхідно проаналізувати показники часу, що було затрачено на кожен етап завантаження даних у БД, а також виміряти швидкість запису даних у БД. Виконаємо 5 ітерацій завантаження даних з ІС "Конкурс" та їх запис у БД на

комп'ютері з наступними параметрами: ОЗП: 6 Гб, ЦП: 4-ядерний AMD A6-3420m, 1,5 ГГц, HDD: 500Гб (Комп'ютер 1), та визначимо середній показник. Результати наведено у таблиці 4.1:

Таблиця 4.1 - Середні показники часу і швидкості завантаження та запису даних до БД на комп'ютері 1

Показник	Середнє значення
Час краулінгу ІС "Конкурс"	10,21 хвилин
Час завантаження сторінок ІС "Конкурс"	22,42 хвилин
Час запису даних до БД	104,62 хвилин
Розмір БД	276 Мб

Завантаження даних у БД відбувається один раз, оскільки дані за 2014 рік на ІС "Конкурс" не оновлюються. Тому швидкість завантаження даних можна вважати прийнятною, оскільки користувачу необхідно очікувати близько двох годин. Середні показники можна покращити, виконавши завантаження на більш потужному комп'ютері. Виконаємо 5 ітерацій завантаження даних з ІС "Конкурс" та запис їх у БД на комп'ютері із наступними параметрами: ОЗП: 8Гб, ЦП: 4-ядерний 4th Gen Intel® Core™ i7-4700MQ, SSD: 1Тб (Комп'ютер 2), та визначимо середній показник. Результати наведено у таблиці 4.2:

Таблиця 4.2 - Середні показники часу і швидкості завантаження та запису даних до БД на комп'ютері 2

Показник	Середнє значення
Час краулінгу ІС "Конкурс"	9,68 хвилин
Час завантаження сторінок ІС "Конкурс"	22,39 хвилин

Час запису даних до БД	37,12 хвилин
Розмір БД	276 Мб

Видно, що більш потужному комп'ютері (Комп'ютер 2), дані завантажуються та записуються до БД швидше, ніж на комп'ютері 1. В цілому, показники продуктивності роботи модуля можна вважати прийнятними.

4.2 Результати проведених експериментів

4.2.1 Експеримент 1: конкуренція ВНЗ за кількістю осіб

У розділі 2 перший експеримент має визначити конкурентів НТУУ "КПІ" за кількість осіб, причому аналіз потрібно виконувати у розрізі загальної кількості вступників (на усі спеціальності) та окремо на технічні спеціальності. Тому експеримент має складатися з наступних кроків:

1. Визначити топ-15 українських ВНЗ за кількістю вступників, що подали документи.
2. Визначити кількість осіб з-поміж вступників до НТУУ "КПІ" та кожного з інших ВНЗ із топ-15, що одночасно подали документи до обох ВНЗ.
3. Визначити кількість спільних вступників відносно до їх загальної кількості у кожному з цих ВНЗ.
4. Із спільних абітурієнтів визначити кількість тих, що подали оригінали до одного чи іншого ВНЗ.

Під час експеримента до вибірки не було включено дані про таких абітурієнтів:

- Вступники, які мають право на позаконкурний вступ;
- Вступники, які подали документи на інші ОКР, крім «Бакалавр».

Для того щоб виконати пункт 1, необхідно виконати наступний запит до новоствореної БД:

```
USE AdmissionDB
```

```
SELECT TOP 15 i.Title, COUNT(DISTINCT en.Id) AS totalEnrollees
```

FROM Enrollments AS et

INNER JOIN Enrollees AS en ON et.EnrolleeId = en.Id

INNER JOIN Institutions AS i ON et.InstitutionId = i.Id

INNER JOIN EnrollmentDirections AS ed ON et.EnrollmentDirectionId = ed.Id

INNER JOIN Degrees AS dg ON ed.DegreeId = dg.Id

WHERE DegreeId = 3 and et.OutOfContest = 0 and i.Title != N'Національний технічний університет України
"Київський політехнічний інститут"'

GROUP BY i.Title

ORDER BY totalEnrollees DESC

Результати виконання запиту наведені у рис. 4.8.

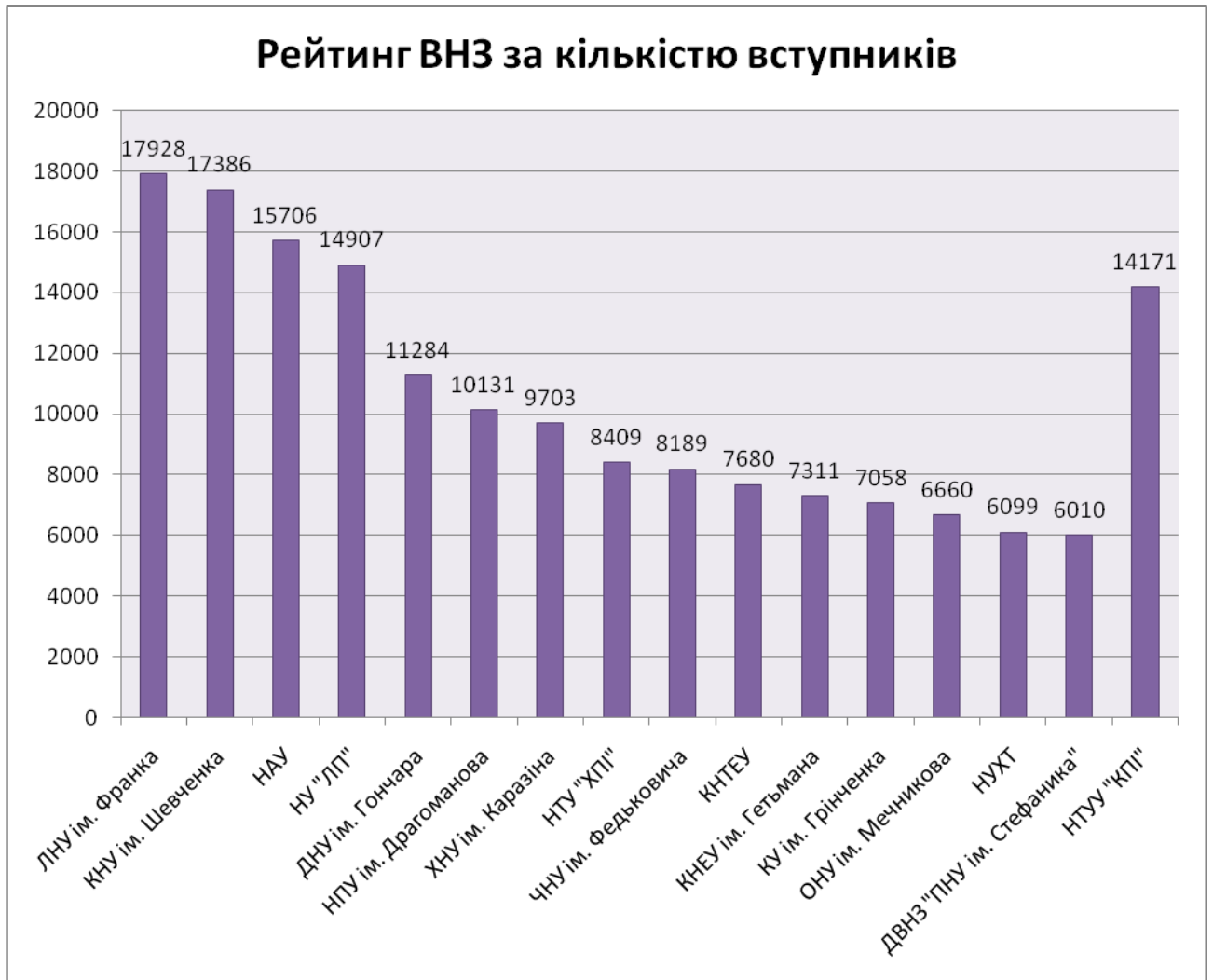


Рисунок 4.8 - Топ-16 ВНЗ України за кількістю вступників

Проведемо підрахунок згідно з пунктами 2-4 експерименту 1 та запишемо результати до таблиці 4.3. Запит, що використовувався для формування статистики, наведено у "Додатку В".

Таблиця 4.3 - Результати попарного порівняння НТУУ "КПІ" з іншими ВНЗ по загальній кількості вступників

ВНЗ	Спільних вступників із НТУУ "КПІ"	% від загальної кількості вступників		Зарахованих	
		НТУУ "КПІ"	Цей ВНЗ	НТУУ "КПІ"	Цей ВНЗ
НАУ	5230	36,9	33,29	2221	719
КНУ ім. Шевченка	5122	36,14	29,46	1940	1079
НУ "ЛП"	2333	16,46	15,65	721	463
КНЕУ ім. Гетьмана	2174	15,34	29,73	830	449
КНТЕУ	1559	11	20,29	481	281
НУХТ	1508	10,64	24,72	581	306
НПУ ім. Драгоманова	1141	8,05	11,26	220	167
ЛНУ ім. Франка	1138	8,03	6,34	199	145
КУ ім. Грінченка	828	5,84	11,73	133	138
НТУ "ХПІ"	655	4,62	7,78	145	124
ДНУ ім. Гончара	616	4,34	5,45	161	132
ХНУ ім. Каразіна	399	2,81	4,11	101	62
ЧНУ ім. Федьковича	316	2,22	3,85	77	61
ОНУ ім. Мечникова	224	1,58	3,36	49	39
ДВНЗ "ПІНУ ім. Стефаніка"	110	0,77	1,83	23	16

- ВНЗ – навчальний заклад, із яким проводиться парне порівняння НТУУ "КПІ".

- Спільних вступників – кількість вступників, які одночасно подали документи до НТУУ "КПІ" та іншого ВНЗ.

- % від загальної кількості вступників – кількість спільних вступників відносно загальної кількості осіб, що подали документи на вступ до НТУУ "КПІ" або до іншого ВНЗ.

- Зарахованих – кількість спільних вступників, що подали оригінали документів до НТУУ "КПІ" або іншого ВНЗ.

Аналогічним чином отримаємо результати експерименту серед вступників на технічні спеціальності. Рейтинг ВНЗ за кількістю вступників на технічні спеціальності та порівняння НТУУ "КПІ" з іншими ВНЗ наведено у рис. 4.9 та таблиці 4.4.



Рисунок 4.9 - Топ-16 ВНЗ України за кількістю вступників на технічні напрями

Таблиця 4.4 - Результати попарного порівняння НТУУ "КПІ" із іншими ВНЗ по кількості вступників на технічні спеціальності

ВНЗ	Спільних вступників із НТУУ "КПІ"	% від загальної кількості вступників		Зарахованих	
		НТУУ "КПІ"	Цей ВНЗ	НТУУ "КПІ"	Цей ВНЗ
КНУ ім. Шевченка	1354	34,54	52,27	672	179
НАУ	1337	34,11	43,05	617	182
НУ "ЛПІ"	990	25,26	28,23	275	196

НУКМА	487	12,42	43,52	243	47
ЛНУ ім. Франка	335	8,54	16,8	54	51
КНЕУ ім. Гетьмана	288	7,34	29,78	156	18
НТУ "ХПІ"	286	7,29	10,71	53	49
ВНТУ	269	6,86	26,55	81	53
КНУБА	238	6,07	28,84	101	29
ХНУРЕ	232	5,91	11,5	80	35
ОНПУ	191	4,87	13,67	37	45
ДНУ ім. Гончара	163	4,15	12,86	34	47
ХНУ ім. Каразіна	101	2,57	7,36	24	12
НАУ ХАІ	95	2,42	7,12	16	19
ДНВЗ "НГУ"	77	1,96	7,18	15	14

З огляду на дані, наведені у таблиці 4.3, можна зробити висновок про те, що найбільш сильними конкурентами НТУУ "КПІ" за кількість вступників є наступні шість ВНЗ:

1. Національний авіаційний університет;
2. Київський національний університет імені Тараса Шевченка;
3. Національний університет "Львівська політехніка";
4. Київський національний економічний університет ім. Вадима Гетьмана;
5. Київський національний торговельно-економічний університет;
6. Національний університет харчових технологій.

Про це свідчить кількість спільник абітурієнтів відносно загальної кількості вступників до ВНЗ (більше десяти відсотків). Але, як видно із наведених даних, у чотирнадцяти із п'ятнадцяти випадків вступники віддають перевагу НТУУ "КПІ".

Якщо відсортувати обрані топ-15 ВНЗ за кількістю спільних вступників, то можна помітити, що найбільша кількість таких абітурієнтів у київських навчальних закладах.

Як видно із таблиці 4.4, у цілому, НТУУ "КПІ" також поза конкуренцією серед вступників на технічні напрями. У тринадцяти із п'ятнадцяти випадків абітурієнти надають перевагу НТУУ "КПІ". Найбільшими конкурентами за технічними спеціальностями є наступні ВНЗ:

1. Національний авіаційний університет;
2. Київський національний університет імені Тараса Шевченка;
3. Національний університет "Львівська політехніка";
4. Національний університет "Києво-Могилянська академія".

Із наведеними ВНЗ НТУУ "КПІ" має найбільшу кількість спільних вступників (більше десяти відсотків). Як і у першій частині експерименту, видно, що більшість ВНЗ, із якими конкурує НТУУ "КПІ", знаходяться у Києві.

4.2.2 Експеримент 2: конкуренція ВНЗ за кількістю поданих заяв

У розділі 2 другий експеримент має визначити конкурентів НТУУ "КПІ" за кількістю поданих заяв. При цьому аналіз потрібно виконувати у розрізі загальної кількості заяв (на усі спеціальності) та окремо на технічні спеціальності. Тому експеримент складається з таких кроків:

1. Визначити топ-15 українських ВНЗ за кількістю поданих заяв.
2. Визначити кількість осіб з-поміж вступників до НТУУ "КПІ" та кожного з інших ВНЗ із топ-15, що одночасно подали документи до обох ВНЗ.
3. Визначити отриману кількість заяв вступників відносно загальної кількості кожного з цих ВНЗ.
4. Із спільних абітурієнтів визначити кількість тих, що подали оригінали до одного чи іншого ВНЗ.

У дослідженні не буде враховано:

- Заяви вступників, які мають право на позаконкурний вступ;

– Заяви вступників, які не подали документи інші ОКР, окрім «Бакалавр».

Для того щоб виконати пункт 1 другого експерименту, необхідно

виконати запит до БД:

```
USE AdmissionDB;
SELECT TOP 15 i.Title, COUNT(DISTINCT et.Id) AS totalTechEnrollees
from Enrollments AS et
INNER JOIN Enrollees AS en ON et.EnrolleeId = en.Id
INNER JOIN Institutions AS i ON et.InstitutionId = i.Id
INNER JOIN EnrollmentDirections AS ed ON et.EnrollmentDirectionId = ed.Id
INNER JOIN Degrees AS dg ON ed.DegreeId = dg.Id
INNER JOIN Directions AS d ON d.Id = ed.DirectionId
WHERE DegreeId = 3
AND et.OutOfContest = 0
AND i.Title != N'Національний технічний університет України "Київський політехнічний інститут"'
GROUP BY i.Title
ORDER BY totalTechEnrollees DESC
```

Результати виконання запиту наведені у рис. 4.10.

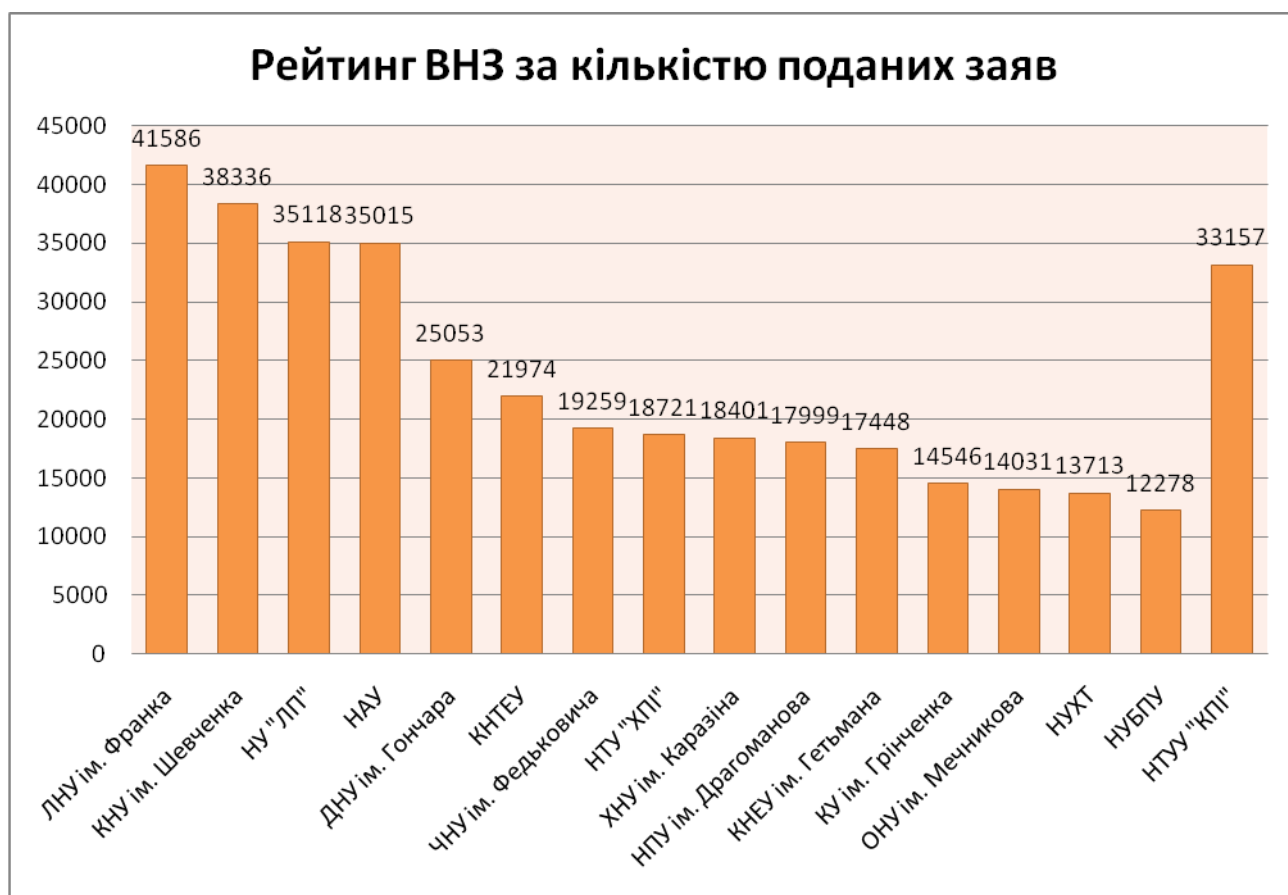


Рисунок 4.10 - Топ-16 ВНЗ України за кількістю поданих вступниками заяв

Проведемо підрахунок згідно з пунктами 2-4 експерименту 1 та запишемо результати до таблиці 4.5. Запит, що використовувався для формування статистики, наведено у "Додатку В".

Таблиця 4.5 - Результати попарного порівняння НТУУ "КПІ" із іншими ВНЗ по загальній кількості заяв вступників

ВНЗ	Кількість поданих заяв спільними вступниками із НТУУ "КПІ"	% від загальної кількості заяв		Зарахованих	
		НТУУ "КПІ"	Цей ВНЗ	НТУУ "КПІ"	Цей ВНЗ
НАУ	26557	80,3	75,84	2185	695

КНУ ім. Шевченка	24523	74,15	63,96	1888	1046
НУ "ЛП"	11577	35	32,96	709	447
КНЕУ ім. Гетьмана	10997	33,25	63,02	817	435
КНТЕУ	8643	26,13	39,33	470	272
НУХТ	7242	21,89	52,81	568	300
ЛНУ ім. Франка	5576	16,86	13,4	195	140
НПУ ім. Драгоманова	4642	14,03	25,79	218	158
НУБПУ	4407	13,32	36,89	264	202
КУ ім. Грінченка	3587	10,84	24,65	128	138
НТУ ХП	3228	9,76	17,24	140	120
ДНУ ім. Гончара	2947	8,91	11,76	158	131
ХНУ ім. Каразіна	1768	5,34	9,6	97	61
ЧНУ ім. Федьковича	1632	4,93	8,47	77	60
ОНУ ім. Мечникова	1010	3,05	7,19	49	38

– ВНЗ – навчальний заклад, із яким проводиться парне порівняння НТУУ "КП".

– Кількість поданих заяв спільних вступників із НТУУ "КП" – кількість поданих заяв вступників, які одночасно подали документи до НТУУ "КП" та іншого ВНЗ.

– % від загальної кількості заяв – кількість заяв спільних вступників відносно загальної кількості поданих заяв.

– Зарахованих – кількість спільних вступників, що подали оригінали документів до НТУУ "КП" або іншого ВНЗ.

Аналогічним чином отримуємо результати експерименту серед заяв вступників на технічні спеціальності. Рейтинг ВНЗ за кількістю заяв вступників на технічні спеціальності та порівняння НТУУ "КП" з іншими ВНЗ наведено у рис. 4.11 та табл. 4.6.

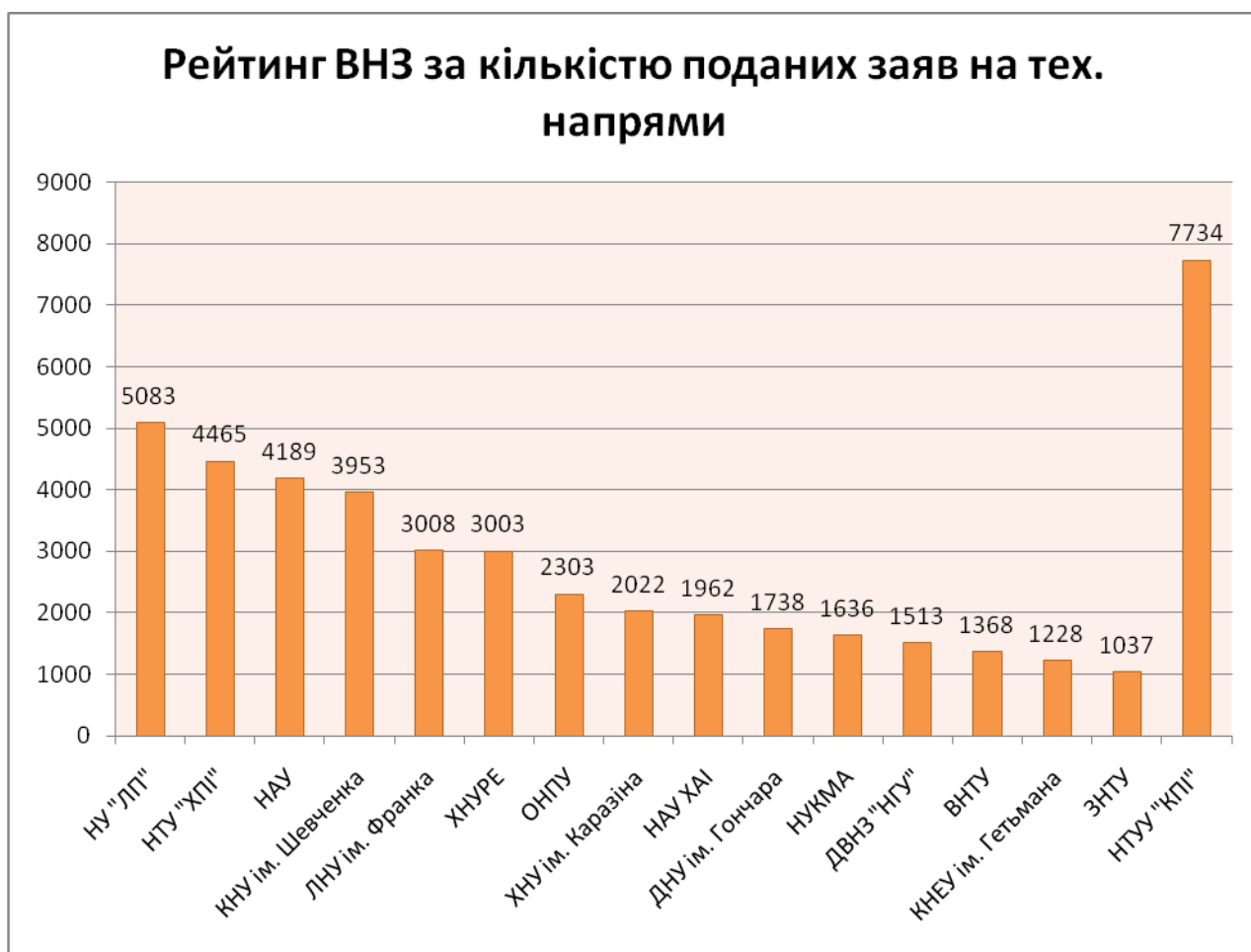


Рисунок 4.11 - Топ-16 ВНЗ України за кількістю абітурієнтів, що подали документи на вступ на технічні спеціальності

Таблиця 4.6 - Результати попарного порівняння НТУУ "КП" із іншими ВНЗ по кількості заяв вступників на технічні спеціальності

ВНЗ	Кількість поданих заяв спільних вступників із НТУУ "КП"	% від загальної кількості заяв		Зарахованих	
		НТУУ "КП"	Цей ВНЗ	НТУУ "КП"	Цей ВНЗ
НУ "ЛП"	26557	80,3	75,84	2185	695
НТУ "ХП"	24523	74,15	63,96	1888	1046
НАУ	11577	35	32,96	709	447

КНУ ім. Шевченка	10997	33,25	63,02	817	435
ЛНУ ім. Франка	8643	26,13	39,33	470	272
ХНУРЕ	7242	21,89	52,81	568	300
ОНПУ	5576	16,86	13,4	195	140
ХНУ ім. Каразіна	4642	14,03	25,79	218	158
НАУ ХАІ	4407	13,32	36,89	264	202
ДНУ ім. Гончара	3587	10,84	24,65	128	138
НУКМА	3228	9,76	17,24	140	120
ДВНЗ "НГУ"	2947	8,91	11,76	158	131
ВНТУ	1768	5,34	9,6	97	61
КНЕУ ім. Гетьмана	1632	4,93	8,47	77	60
ЗНТУ	1010	3,05	7,19	49	38

4.2.3 Експеримент 3: Визначення ВНЗ-конкурента за допомогою методів кластерного аналізу

Як згадувалось у розділі 2, визначити конкурента також можна за допомогою методів кластерного аналізу. Алгоритм проведення експерименту:

1. У якості системи змінних, з одного боку, буде обрано кількість вступників, що одночасно подали документи на вступ до НТУУ "КПІ" та іншого серед топ-15 ВНЗ України за кількістю вступників, визначеному в першому експерименті. З іншого – кількість спільних напрямів підготовки НТУУ "КПІ" та іншого ВНЗ.

2. Щоб уникнути «домінування» змінних з великим масштабом виміру, буде проведено попереднє нормування вихідних змінних. Для спільних вступників буде визначена відносна кількість осіб до загальної кількості абітурієнтів НТУУ "КПІ", та аналогічно для напрямів – буде визначено

відносну кількість спільних бакалавратів до загальної кількості напрямів НТУУ "КПІ".

3. Буде проведено розбиття на групи, причому, належність об'єкта до групи буде визначатися за допомогою формули 2.1.

Отже, для виконання першого кроку експерименту, по-перше, необхідно визначити загальну кількість напрямів підготовки серед топ-15 ВНЗ за кількістю вступників. Для цього потрібно виконати запит до створенної БД:

```
USE AdmissionDB
SELECT i.Title, COUNT(DISTINCT d.Id) AS NumberOfDirections
FROM Enrollments AS et
INNER JOIN Institutions AS i ON et.InstitutionId = i.Id
INNER JOIN EnrollmentDirections AS ed ON et.EnrollmentDirectionId = ed.Id
INNER JOIN Degrees AS dg ON ed.DegreeId = dg.Id
INNER JOIN Directions AS d ON d.Id = ed.DirectionId
WHERE DegreeId = 3 AND i.Title IN(
N'Національний технічний університет України "Київський політехнічний інститут"',
N'Львівський національний університет імені Івана Франка',
N'Київський національний університет імені Тараса Шевченка',
N'Національний авіаційний університет',
N'Національний університет "Львівська політехніка"',
N'Дніпропетровський національний університет імені Олеся Гончара',
N'Національний педагогічний університет імені М.П. Драгоманова',
N'Харківський національний університет імені В.Н. Каразіна',
N'Національний технічний університет "Харківський політехнічний інститут"',
N'Чернівецький національний університет імені Юрія Федьковича',
N'Київський національний торговельно-економічний університет',
N'Київський національний економічний університет імені Вадима Гетьмана, ДВНЗ',
N'Київський університет імені Бориса Грінченка',
N'Одеський національний університет імені І.І. Мечникова',
N'Національний університет харчових технологій',
N'Державний вищий навчальний заклад "Прикарпатський національний університет імені Василя Стефаника"
)
AND et.OutOfContest = 0
GROUP BY i.Title
ORDER BY NumberOfDirections DESC
```

Результати виконання запиту наведено у рис. 4.12:

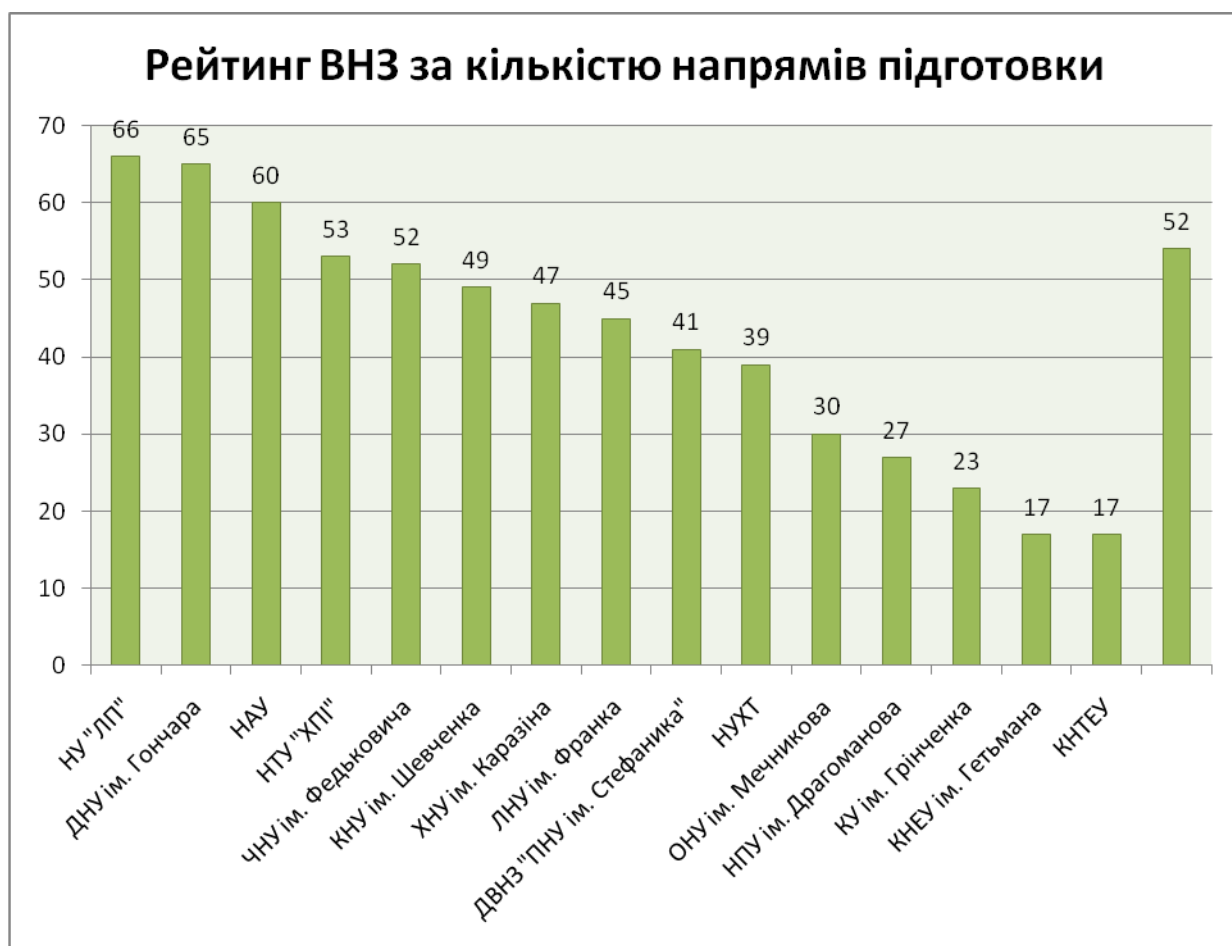


Рисунок 4.12 - Рейтинг ВНЗ за кількістю напрямів підготовки студентів

Наступним кроком є визначення кількості спільних вступників та напрямів підготовки, а також подальше нормування показників. Кількість спільних вступників до НТУУ "ХП" та інших серед топ-15 ВНЗ було визначено в експерименті 1, тому можна скористатись отриманими раніше даними. Для визначення кількості спільних напрямів необхідно виконати запит до створеної БД та виконати нормування. Результати наведено у табл. 4.7

Таблиця 4.7 - Результати попарного порівняння НТУУ "КПІ" із іншими ВНЗ за кількістю спільних напрямів підготовки

ВНЗ	Кількість спільних напрямів підготовки	% спільних напрямів підготовки	
		НТУУ "КПІ"	Цей ВНЗ
НУ "ЛП"	40	74,07	60,6
ДНУ ім. Гончара	34	56,66	62,96
НАУ	34	62,96	64,15
НТУ "ХПІ"	33	61,11	70,21
ЧНУ ім. Федьковича	30	55,55	46,15
КНУ ім. Шевченка	22	40,74	42,3
ХНУ ім. Каразіна	19	35,18	42,22
ЛНУ ім. Франка	16	29,62	32,65
ДВНЗ "ПНУ ім. Стефаника"	15	27,77	36,58
НУХТ	15	27,77	65,21
ОНУ ім. Мечникова	13	24,07	43,33
НПУ ім. Драгоманова	12	22,22	30,76
КУ ім. Грінченка	9	16,66	33,33
КНЕУ ім. Гетьмана	9	16,66	52,94
КНТЕУ	9	16,66	52,94

Третій крок алгоритму можна виконази за допомогою спеціалізованого програмного забезпечення. Для проведення цього експерименту буде використаний пакет Wolfram Mathematica. Це система комп'ютерної алгебри

компанії Wolfram Research, що містить багато функцій як для аналітичних перетворень, так і для чисельних розрахунків. Крім того, програма підтримує роботу з графікою і звуком, включаючи побудову дво- і тривимірних графіків функцій, малювання довільних геометричних фігур, імпорт та експорт зображень і звуку[10].

Серед досить великого набору функцій чисельних розрахунків, система також дає гнучкі інструменти для проведення кластерного аналізу. За допомогою них можна виконувати групування об'єктів будь-яким із наведених у другому розділі методів групування даних. Також програмний пакет надає широкі можливості представлення вихідних даних, що дозволяє швидко та чітко зробити висновки.

Отже, для визначення кластерів, необхідно виконати наступний скрипт:

```
datarecords1 = {"НУ ЛП", 74.07, 16.46}, {"ЧНУ ім. Федьковича", 55.55, 2.22},
  {"ДНУ ім. Гончара", 56.66, 4.34}, {"НАУ", 62.96, 36.9}, {"КНУ ім. Шевченка", 40.74, 36.14},
  {"ЛНУ ім. Франка", 29.62, 8.03}, {"НТУ ХПІ", 61.11, 4.62}, {"ХНУ ім. Каразіна", 35.18, 2.81},
  {"ДВНЗ ПНУ ім. Стефаника", 27.77, 0.77}, {"НПУ ім. Драгоманова", 22.22, 8.05},
  {"ОНУ ім. Мечникова", 24.07, 3.36}, {"КУ ім. Грінченка", 16.66, 5.84}, {"НУХТ", 27.77, 10.64},
  {"КНЕУ ім. Гетьмана", 16.66, 15.34}, {"КНТЕУ", 16.66, 11}};
c1 = FindClusters[datarecords1, DistanceFunction -> EuclideanDistance]
ListPlot[Tooltip[datarecords1], PlotStyle -> PointSize -> Large, AxesLabel -> {Directions, Enrollees},
  AspectRatio -> Automatic, AxesOrigin -> {0, 0}]
```

Результати зображено на рис. 4.13.

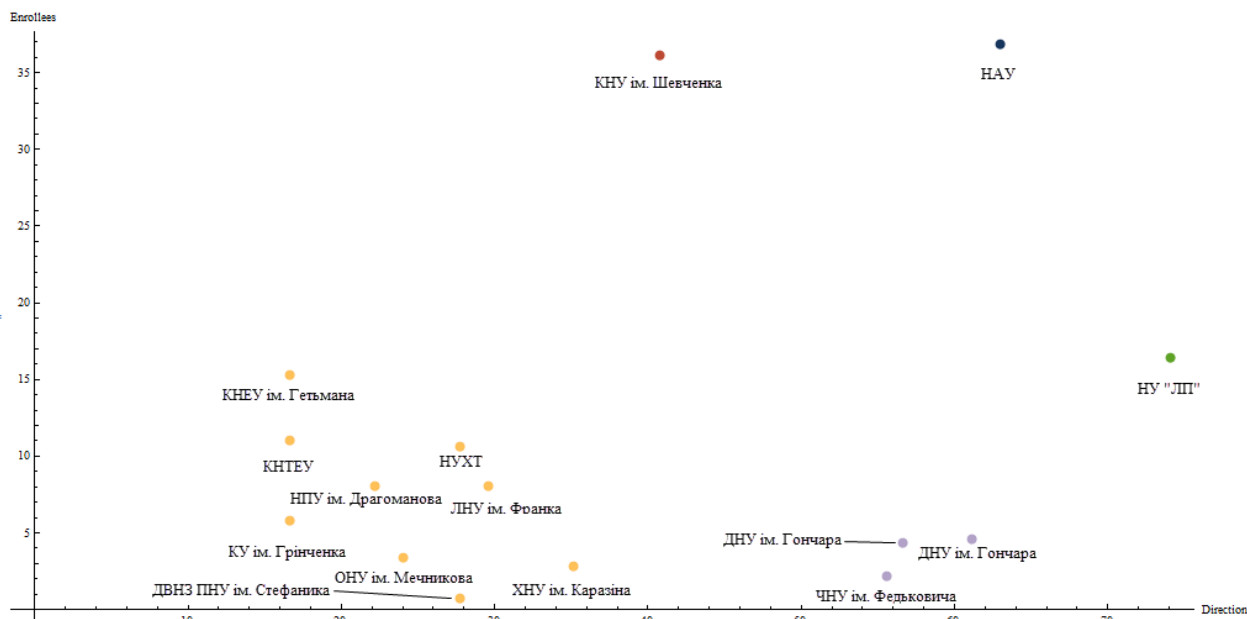


РИСУНОК 4.13 - РОЗПОДІЛ ОБ'ЄКТІВ НА КЛАСТЕРИ

Як можна помітити, за допомогою Wolfram Mathematica та методу найближчого сусіда вдалося виділити 5 кластерів (на рис. 4.13 вони зображені різними кольорами). Проаналізуємо отримані групи об'єктів.

1. Жовта група. До групи увійшло 9 ВНЗ. Спільним для цих навчальних закладів можна вважати те, що вони мають досить невелику відносну кількість спільних із НТУУ "КПІ" вступників, а також невелику кількість співпадінь серед напрямів підготовки студентів. Можна зробити висновок про те, що ця група ВНЗ не є конкурентами для НТУУ "КПІ".

2. Фіолетова група. До групи увійшло 3 ВНЗ. Особливістю ВНЗ цієї групи є те, що вони мають досить велику кількість спільних напрямів підготовки із НТУУ "КПІ". Водночас, кількість спільних вступників досить незначна. Навчальні заклади цієї групи також можна не вважати конкурентами для НТУУ "КПІ", але варто зазначити, що вони мають потенціал стати ними, оскільки кількість спільних напрямів досить велика.

3. Зелена група. До групи увійшов один ВНЗ. Національний університет "Львівська політехніка" можна вважати одним із конкурентів НТУУ "КПІ". Серед досліджуваних ВНЗ НУ "ЛП" має найбільшу відносну кількість спільних напрямів із НТУУ "КПІ", хоча і відносна кількість спільних вступників не

найбільша з-поміж навчальних закладів, що розглядаються. Цей ВНЗ має найбільший потенціал та має всі шанси стати досить сильним конкурентом НТУУ "КПІ" у майбутньому.

4. Червона група. До групи увійшов один ВНЗ - Київський національний університет імені Тараса Шевченка. Характеристика цієї групи зворотня до зеленої групи: кількість спільних напрямів не найбільша, але відносна кількість спільних вступників є однією з найбільших. Така велика кількість спільних абітурієнтів може бути спричинено тим фактом, що КНУ ім. Шевченка є одним із найвідоміших ВНЗ країни, а отже на вибір вступника може впливати саме престижність цього закладу. Цей ВНЗ - другим сильним конкурентом НТУУ "КПІ".

5. Синя група. До групи увійшов один ВНЗ - НАУ. Національний авіаційний університет у 2014 році є найбільшим конкурентом НТУУ "КПІ". Цей навчальний заклад має найбільшу відносну кількість спільних вступників. Таку велику кількість спільних вступників можна пояснити тим, що ці два ВНЗ мають досить багато спільних напрямів підготовки, а отже під час вступної кампанії абітурієнти подають документи до двох ВНЗ із один із них використовується в якості резервного варіанту. Також на це може впливати те, що ці заклади одними із найпрестижніших в Україні та територіально вони розташовані досить близько.

Як і раніше, у роботі експерименти проводяться для загальної кількості вступників і для вступників на технічні напрями. У якості досліджуваних напрямів оберем усі, які пов'язані із інформаційними технологіями. Для того, щоб сформувати рейтинг ВНЗ серед вступників на ІТ-напрями необхідно виконати наступний запит до створенної БД:

```
USE AdmissionDB
```

```
SELECT TOP 16 i.Title, COUNT(DISTINCT en.Id) AS totalEnrollees
```

```
FROM Enrollments AS et
```

```
INNER JOIN Enrollees AS en ON et.EnrolleeId = en.Id
```

```
INNER JOIN Institutions AS i ON et.InstitutionId = i.Id
```

```
INNER JOIN EnrollmentDirections AS ed ON et.EnrollmentDirectionId = ed.Id
```

```

INNER JOIN Directions AS d ON d.Id = ed.DirectionId
INNER JOIN Degrees AS dg ON ed.DegreeId = dg.Id
INNER JOIN Branches AS b ON b.Id = ed.BranchId
WHERE DegreeId = 3 AND et.OutOfContest = 0
AND BranchId IN (16, 18, 14, 27, 13)
GROUP BY i.Title
ORDER BY totalEnrollees DESC

```

Результати виконання запиту наведено у рис. 4.14.

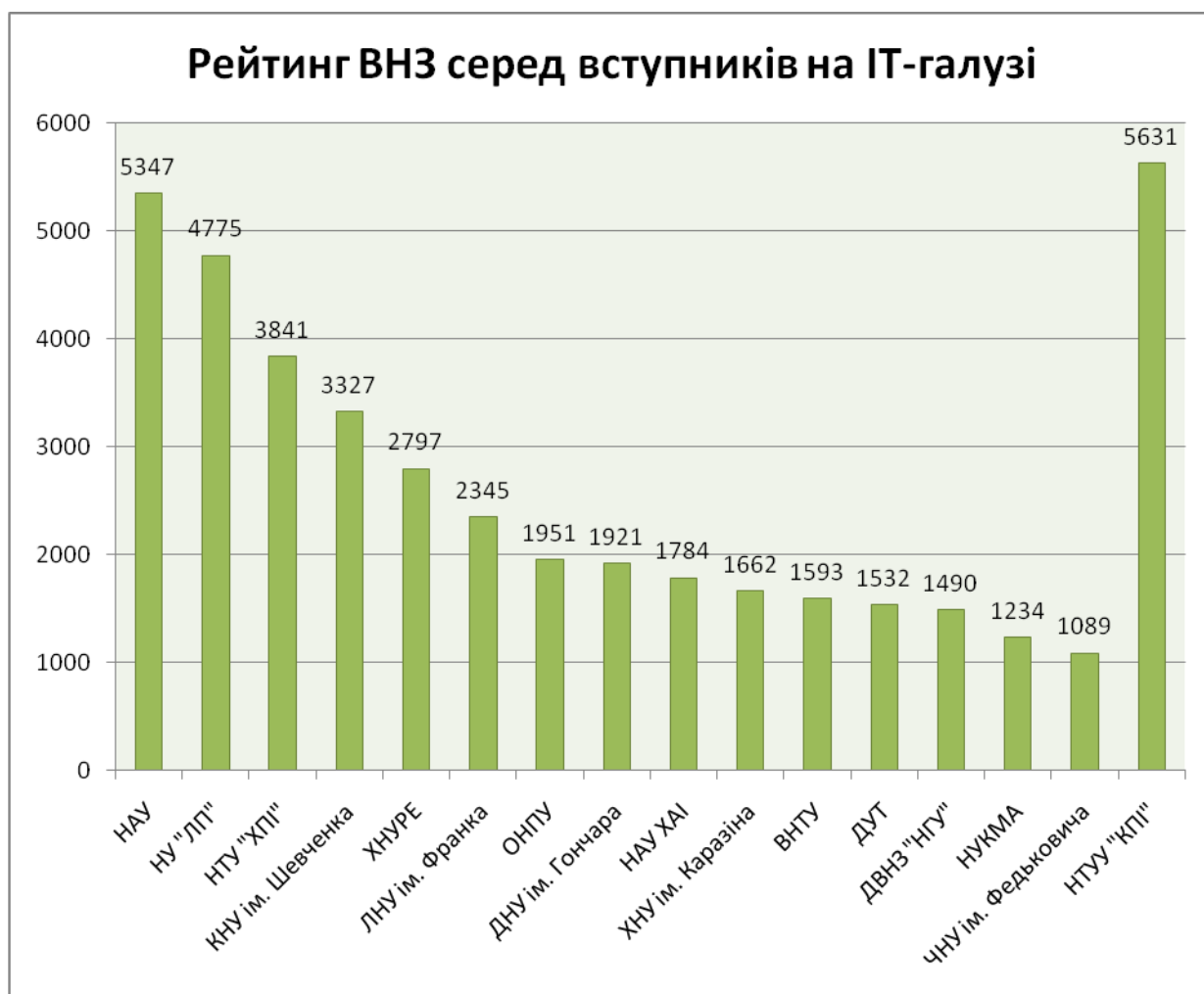


Рисунок 4.14 - Топ-15 ВНЗ за кількістю вступників на ІТ-галузі

Наступним кроком експерименту є визначення кількості спільних вступників та нормування цього параметру, тобто визначення відносної кількості абітурієнтів до загальної кількості. Результати експерименту наведено у табл. 4.8.

Таблиця 4.8 - Результати попарного порівняння НТУУ "КПІ" із іншими ВНЗ за кількістю спільних вступників на ІТ-галузі

ВНЗ	Кількість поданих заяв спільних вступників із НТУУ "КПІ"	% від загальної кількості вступник	
		НТУУ "КПІ"	Цей ВНЗ
НАУ	2240	39,77	41,89
КНУ ім. Шевченко	1868	33,17	56,14
НУ "ЛП"	1391	24,7	29,13
НУКМА	606	10,76	49,1
ДУТ	552	9,8	36,03
ЛНУ ім. Франка	439	7,79	18,72
ВНТУ	380	6,74	23,85
ХНУРЕ	292	5,18	10,43
ОНПУ	269	4,77	13,78
ДНУ ім. Гончара	253	4,49	13,17
НАУ ХАІ	150	2,66	8,4
НТУ "ХПІ"	147	2,61	8,84
ЧНУ ім. Федьковича	136	2,41	12,48
ДВНЗ "НГУ"	113	2	7,58

Іншим критерієм, за яким буде відбуватись групування об'єктів є кількість спільних ІТ-напрямів. За цим параметром також необхідно визначити спільну кількість напрямів НТУУ "КПІ" із іншими досліджуваними ВНЗ та виконати нормування цього параметру. Результати виконання операцій наведено у табл. 4.9.

Таблиця 4.9 - Результати попарного порівняння НТУУ "КПІ" із іншими ВНЗ за кількістю спільних напрямів у ІТ-галузі

ВНЗ	Кількість спільних ІТ-напрямів	% від загальної кількості ІТ-напрямів	
		НТУУ "КПІ"	Цей ВНЗ
НАУ	10	83,33	90,9
КНУ ім. Шевченко	11	84,61	91,66
НУ "ЛП"	10	83,33	90,9
НУКМА	2	16,66	66,66
ДУТ	4	33,33	80
ЛНУ ім. Франка	4	33,33	80
ВНТУ	8	66,66	88,88
ХНУРЕ	12	100	85,71
ОНПУ	8	66,66	88,88
ДНУ ім. Гончара	7	58,33	87,5
НАУ ХАІ	8	66,66	88,88
НТУ "ХПІ"	10	83,33	90,9
ХНУ ім. Каразіна	5	41,66	83,33
ЧНУ ім. Федьковича	8	66,66	88,88
ДВНЗ "НГУ"	8	66,66	88,88

За допомогою пакета Wolfram Mathematica виконаємо побудову кластерів. Для цього необхідно виконати наступний скрипт:

```

datarecords2 = {{39.77, 83.33}, {33.17, 84.61}, {24.7, 83.33}, {10.76, 16.66},
  {9.8, 33.33}, {7.79, 33.33}, {6.74, 66.66}, {5.18, 100}, {4.77, 66.66},
  {4.49, 58.33}, {2.66, 66.66}, {2.61, 83.33}, {2.51, 41.66}, {2.41, 66.66},
  {2, 66.66}};
c2 = FindClusters[datarecords2, DistanceFunction -> EuclideanDistance]
ListPlot[Tooltip[datarecords2], PlotStyle -> PointSize -> Large,
  AxesLabel -> {Enrollees, Directions}, AspectRatio -> Automatic, AxesOrigin -> {0, 0}]

```

Результат виконання скрипта зображено на рис. 4.15.

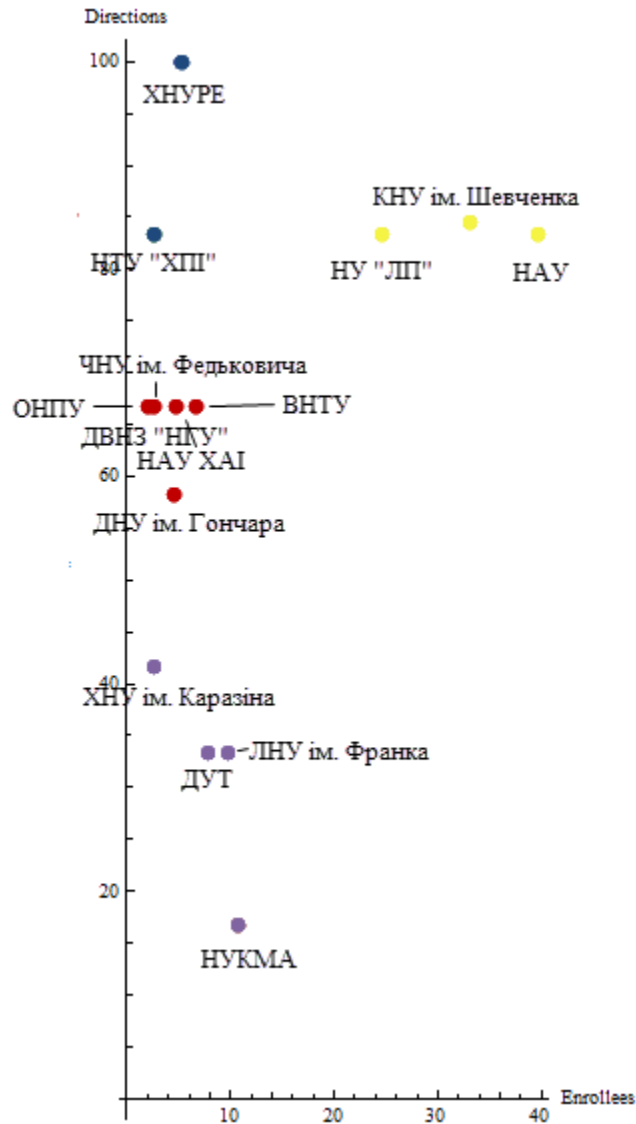


Рисунок 4.15 - Розподіл об'єктів на кластери

1. Фіолетова група. До групи увійшло 4 ВНЗ. Спільним для цих навчальних закладів можна вважати те, що вони мають середню відносну кількість спільних із НТУУ "КПІ" вступників, а також невелику кількість

співпадінь серед напрямів підготовки студентів у ІТ-галузі. Ці ВНЗ можуть вважатися конкурентами для НТУУ "КПІ" тільки за окремими напрямами, а не у цілому.

2. Червона група. До групи увійшло 6 ВНЗ. Єдина відмінність цієї групи від попередньої є те, що ВНЗ мають дещо більшу кількість спільних напрямів із НТУУ "КПІ", але кількість вступників, за яке ці навчальні заклади конкурують така ж незначна. Ця група ВНЗ також не є конкурентом НТУУ "КПІ", хоча і потенціал трохи більший, ніж у фіолетової групи.

3. Синя група. До групи увійшло 2 ВНЗ - ХНУРЕ та НТУ "ХПІ". Заклад синьої групи мають найбільший потенціал серед усіх досліджуваних ВНЗ. Хоча кількість спільних вступників незначна, ці навчальні заклади можуть бути досить сильними конкурентами для НТУУ "КПІ". Також така низька кількість спільних вступників може свідчити про те, що більшість харківських студентів вступає до цих ВНЗ, причому не роздивляються варіант вступу до НТУУ "КПІ". Цей факт може підтвердити рейтинг, зображений на рис. 4.14, де видно, що ці два ВНЗ знаходяться у першій п'ятірці.

4. Жовта група. До групи увійшли 3 ВНЗ - КНУ ім. Шевченка, НУ "ЛПІ" та НАУ. Заклади цієї групи є найбільш сильними конкурентами для НТУУ "КПІ", оскільки кількість спільних напрямів та вступників досить велика, а отже вірогідність того, що абітурієнт віддасть перевагу одному із цих ВНЗ під час вступної кампанії досить велика.

4.2.4 Експеримент 4: Визначення ВНЗ-конкурентів за сильних вступників

У попередніх підрозділах експериментально було визначено трьох найбільших конкурентів НТУУ "КПІ" - КНУ ім. Шевченка, НАУ та НУ "ЛПІ". Раніше згадувалося, що ВНЗ сперечаються не тільки за кількість вступників, а й за тих, що мають високі бали ЗНО та атестату про середню освіту. Отже,

виникає необхідність перевірити, яким з досліджуваних ВНЗ надають перевагу саме сильні абітурієнти.

Для того, щоб визначити це, потрібно провести наступний експеримент:

1. Визначити кількість спільних вступників, що вступили до того чи іншого ВНЗ.

2. Визначити проміжки сум балів сертифікатів ЗНО та балу атестата про середню освіту. Така комбінація обрана тому, що згідно з [1] кожен ВНЗ має право встановлювати додаткові (бонусні) бали для вступників. А оскільки кожен заклад робить це самостійно, то і максимальний конкурсний бал також може бути різним у двох ВНЗ.

3. Визначити кількість студентів першого курсу що мають ту чи іншу суму балів та визначити відносн кількість таких студентів загальній кількості студентів та зобразити це на графіку.

4. Зробити висновки.

Отже, перше, що необхідно зробити, це - визначити спільних кількість абітурієнтів, що стали студентами першого курсу НТУУ "КПІ" чи іншого ВНЗ. Для цього скористаємось результатами першого експерименту, що наведено у табл. 4.3.

Другим кроком необхідно визначити проміжки сум балу атестата про середню освіту та сертифікатів ЗНО. Максимальний бал ЗНО, який може отримати вступник - 200 балів. Для того, щоб абітурієнт мав можливість брати участь у конкурсному відборі, йому необхідно здавати ЗНО з трьох предметів, а отже максимальна сума балів сертифікатів ЗНО - 600 балів. Згідно з [1] бал атестату про середню освіту може бути 60 балів. Отже, у загальному випадку максимальна сума балів сертифікату ЗНО та атестату про середню освіту може бути 660. Враховуючи це, інтервали суми балів будуть наступними:

- більше 650;
- від 600 до 649,5;
- від 550 до 599,5;
- від 500 до 549,5;

- менше 499,5.

Наступне, що необхідно зробити - визначити кількість студентів першого курсу досліджуваних ВНЗ, що отримали ту чи іншу суму балів. Після того, як буде отримано таку кількість студентів, необхідно визначити відношення цієї кількості до загальної кількості абітурієнтів, що вступили до навчального закладу. Щоб отримати таку статистику, необхідно виконати запит до створеної БД. Лістинг запиту наведено у "Додатку В". Результати зображено на рис. 4.16.

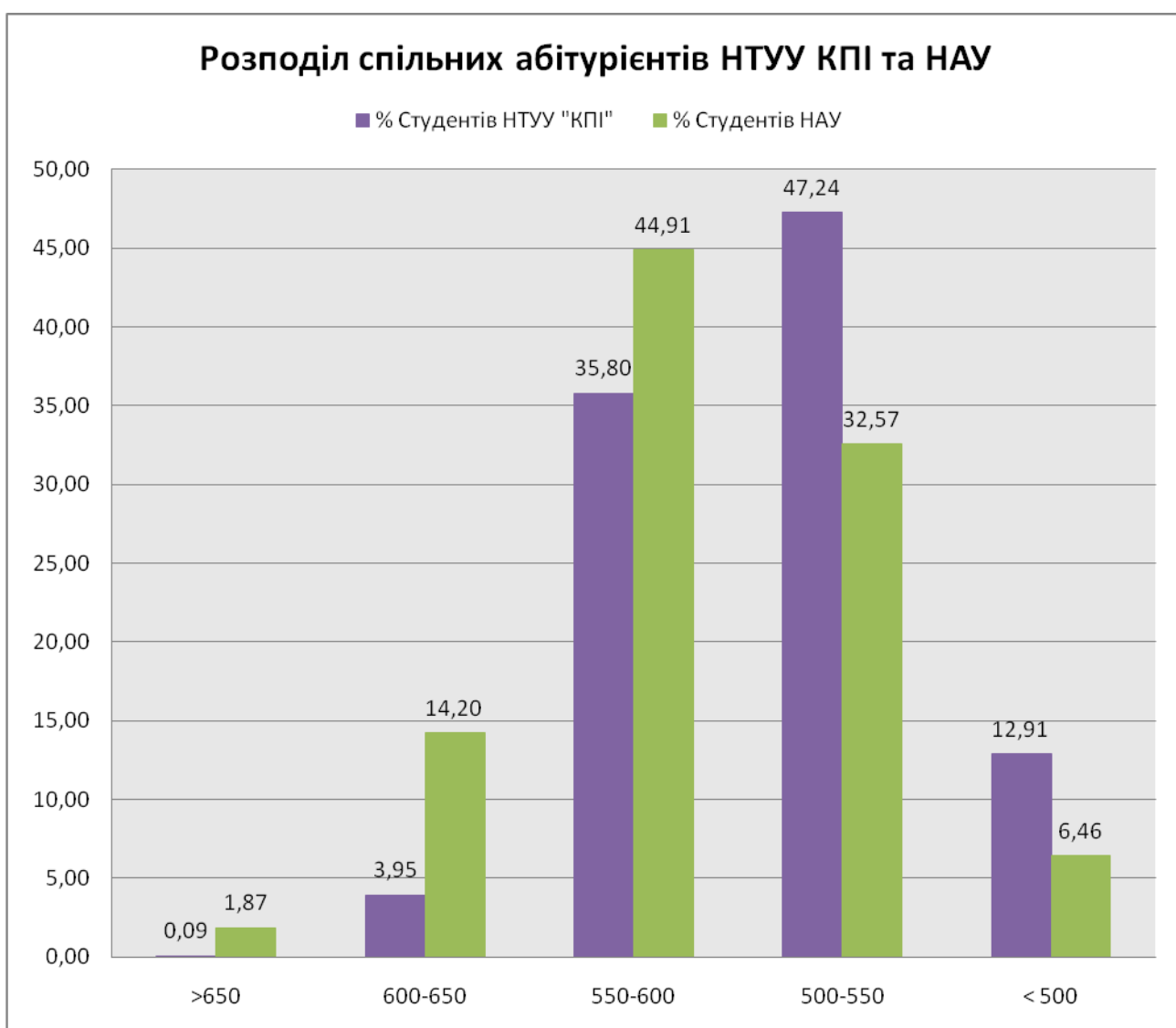


Рисунок 4.16а - Розподіл спільних абітурієнтів, що вступили до НТУУ "КПІ" та НАУ по сумі балів ЗНО та атестату про середню освіту

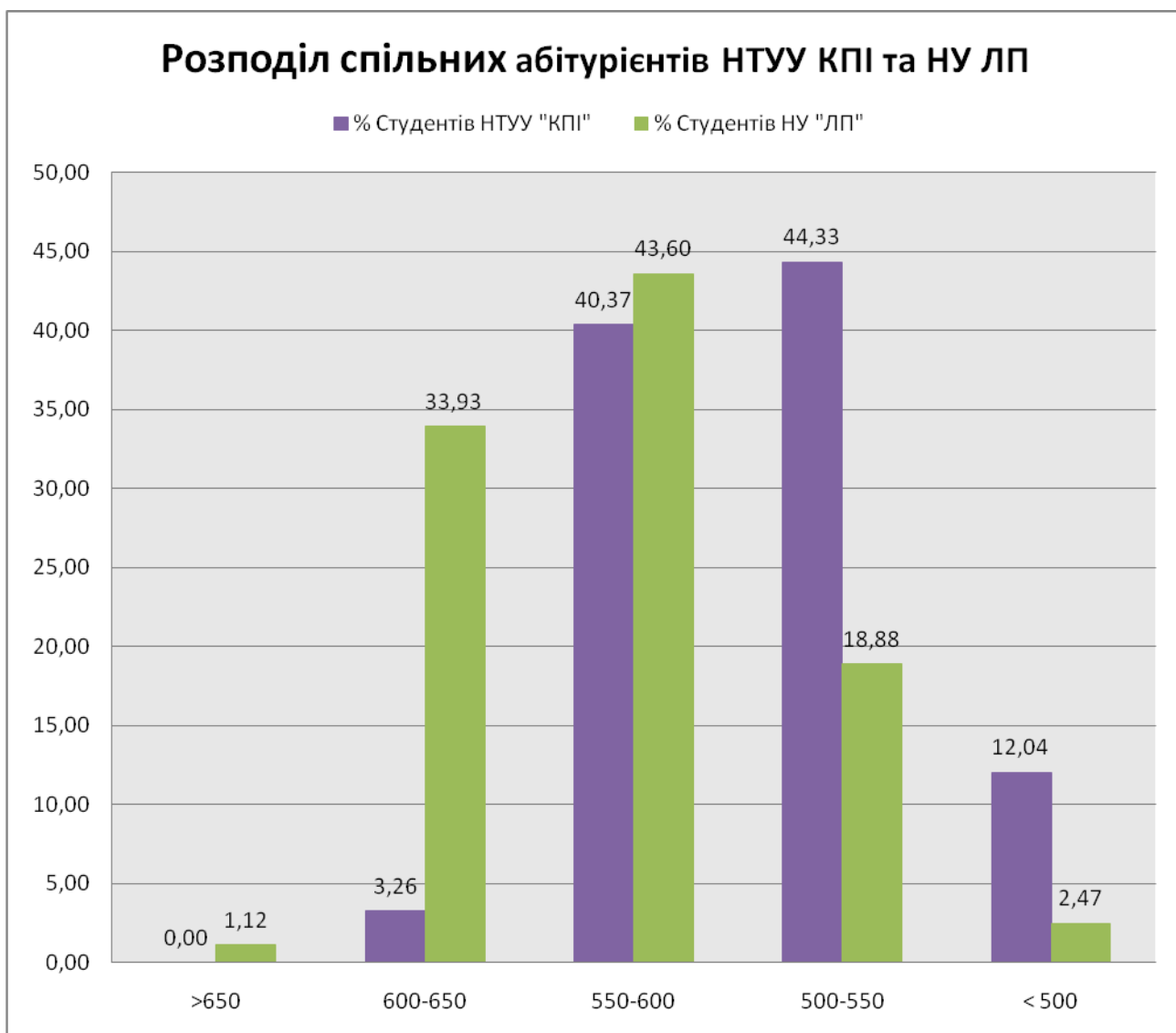


Рисунок 4.16б - Розподіл спільних абітурієнтів, що вступили до НТУУ "КПІ" та НУ "ЛП" по сумі балів ЗНО та атестату про середню освіту

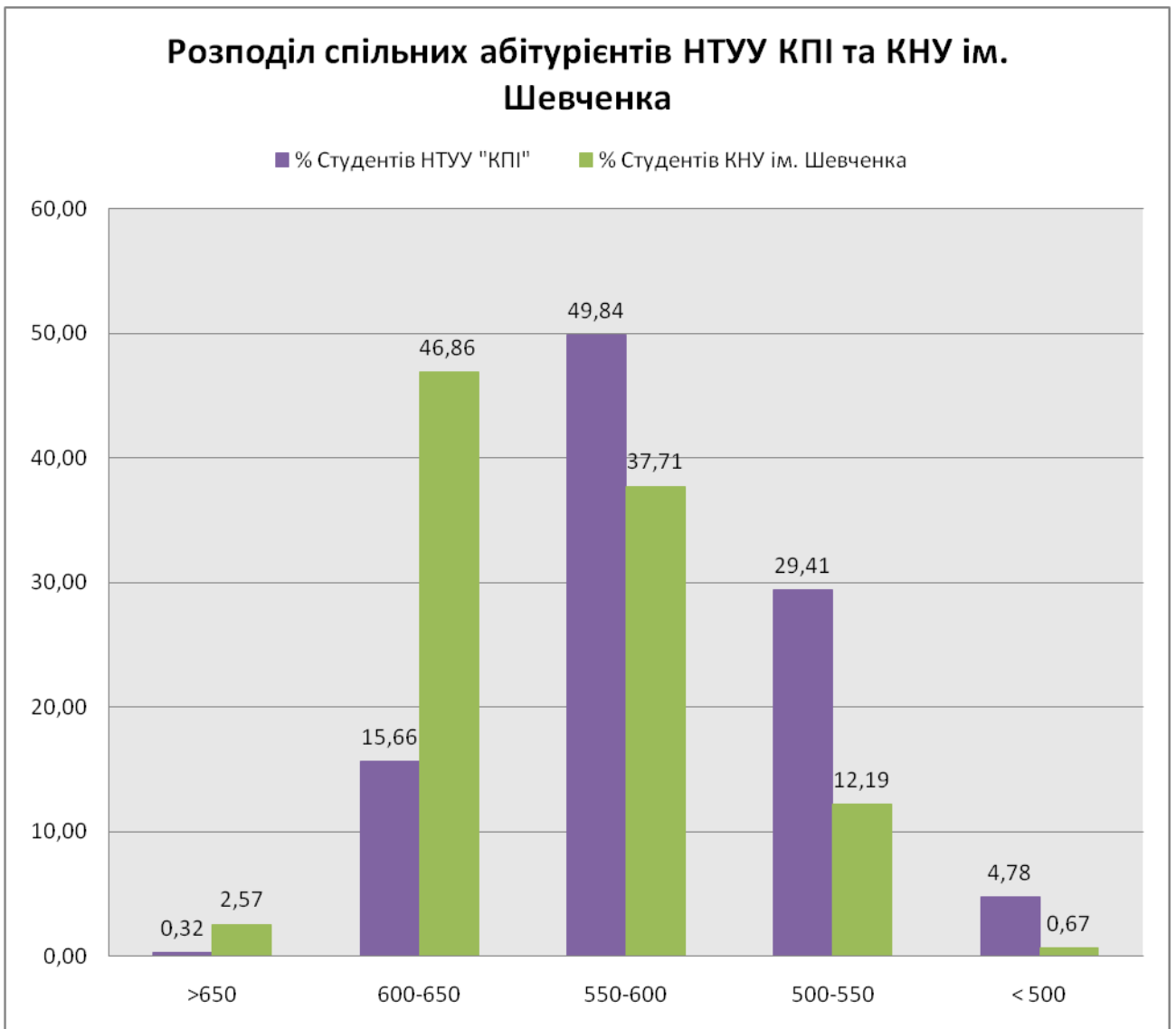


Рисунок 4.16в - Розподіл спільних абітурієнтів, що вступили до НТУУ "КПІ" та КНУ ім. Шевченка по сумі балів ЗНО та атестату про середню освіту

Оскільки кожен експеримент проводиться як для загальної кількості вступників так і для вступників на технічні напрями, необхідно виконати таку ж операцію для вступників на технічні спеціальності. Як і у попередньому експерименті, аналіз буде проводитись серед вступників на усі напрями ІТ-галузей. Для цього необхідно виконати запит до створеної БД. Результати зображено на рис. 4.17.

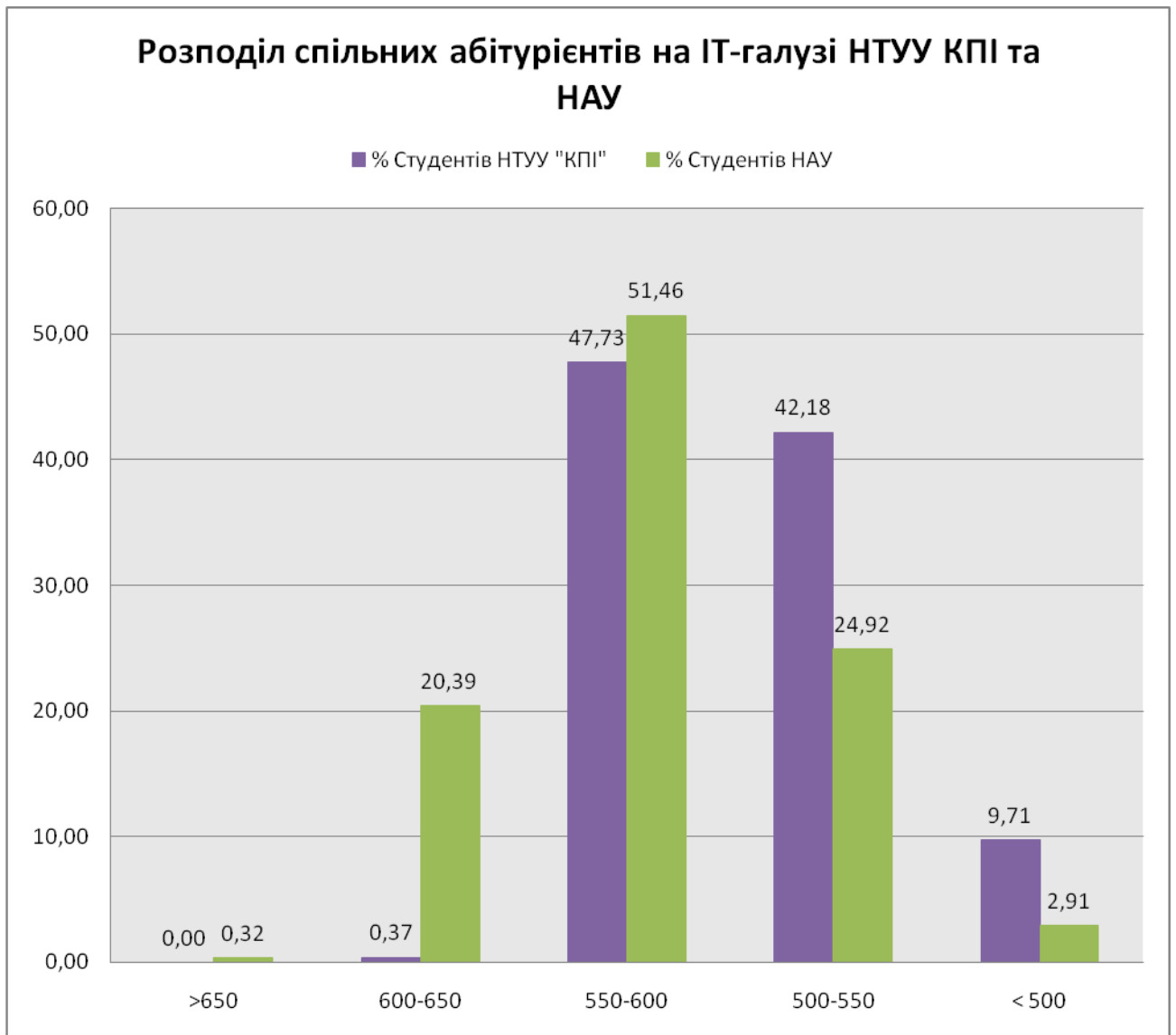


Рисунок 4.17а - Розподіл спільних абітурієнтів на ІТ-галузі, що вступили до НТУУ "КПІ" та НАУ по сумі балів ЗНО та атестату про середню освіту

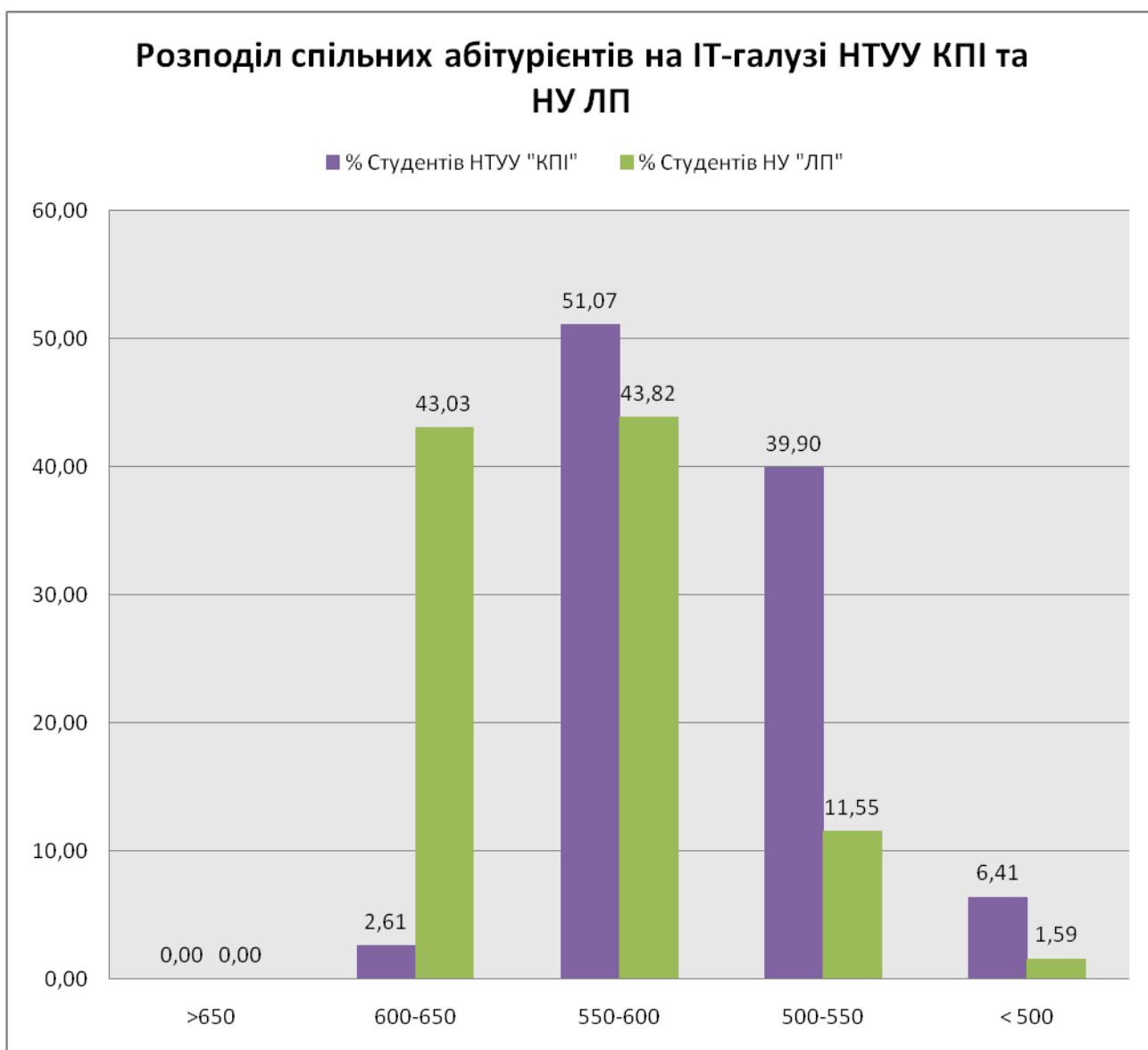


Рисунок 4.17б - Розподіл спільних абітурієнтів на ІТ-галузі, що вступили до НТУУ "КПІ" та НУ "ЛП" по сумі балів ЗНО та атестату про середню освіту

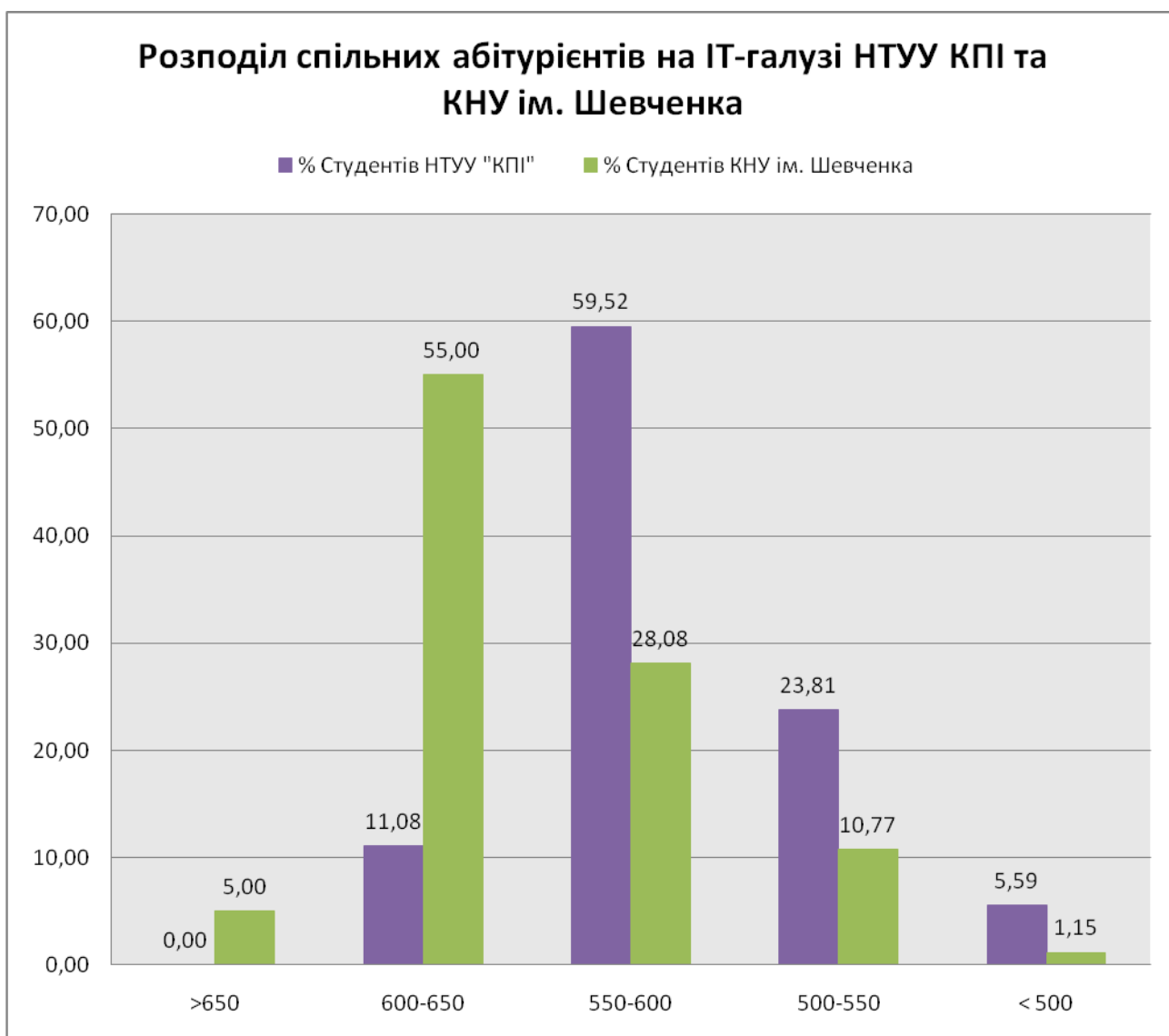


Рисунок 4.17в - Розподіл спільних абітурієнтів на ІТ-галузі, що вступили до НТУУ "КПІ" та КНУ ім. Шевченка по сумі балів ЗНО та атестату про середню освіту

Отже, дивлячись на отримані гістограми можна зробити наступний висновок: не зважаючи на те, що НТУУ "КПІ" обирає більше вступників у порівнянні із найбільшими його конкурентами, розподіл абітурієнтів по сумі балів ЗНО та балу атестату про середню освіту говорить про те, що конкуренти мають більшу відносну кількість сильних вступників. На рис. 4.16а, 4.16б та 4.16в видно, що кількість вступників, які мають суму балів 600-649,5 більше обирали ВНЗ-конкурент ніж НТУУ "КПІ". Із Зниженням суми балів, НТУУ

"КПІ" переважає свого конкурента. Серед вступників на ІТ-галузі аналогічна ситуація, як видно на рис. 4.17а, 4.17б та 4.17в.

4.3 Висновки до четвертого розділу

У даному розділі виконане тестування розробленої системи на швидкість обробки джерела інформації та її завантаження до раніше створеної БД. Результати тестування показали, що швидкість роботи БД є прийнятною для динамічної роботи співробітників приймальної комісії, а також дані, що записуються є повними та не мають зайвої інформації. У даному розділі було проведено чотири експерименти, за допомогою яких можна визначити конкурентів НТУУ "КПІ" у різних розрізах. Причому конкуренти визначались як у випадку загальної кількості вступників так і серед вступників на технічні напрями. Також був проведений експеримент, що використовує методи кластерного аналізу для дослідження. Результати цього експерименту підтвердив результати попередніх експериментів. Отримані результати проаналізовано та були зроблені висновки.

5 ОХОРОНА ПРАЦІ ТА БЕЗПЕКА У НАДЗВИЧАЙНИХ СИТУАЦІЯХ

5.1 Вступ

Темою магістерської дисертації є «Кластеризація переміщень абітурієнтів при поданні документів до ВНЗ». Робота носить дослідницький характер з використанням ресурсів ЕОМ. В результаті роботи було розроблено модуль, що формує базу даних абітурієнтів із інформаційної системи ІС "Конкурс" на мові С# для подальшого формування статистики та її аналізу.

Завданням розділу є аналіз проектованого об'єкту, або виробничих умов, або технологічних процесів, або робочих місць з метою виявлення можливих причин травматизму, профзахворювань, перенавантаження людини.

5.2 Опис приміщення

Згідно законодавства України про умови праці в робочих приміщеннях повинні відповідати нормам охорони праці. Робочий день триває 8 годин.

Так як робота з модулем буде проходити в приміщенні з використання ЕОМ, робоче приміщення має нормуватися згідно з [11], згідно з яким зазначені наступні вимоги до робочого місця

Об'ємно-планувальні рішення будівель та приміщень для роботи з ЕОМ мають задовольняти вимоги [11]

Площа на одне робоче місце має становити не менше ніж, 6,0 м², а об'єм не менше ніж 20,0 м³

Висота робочого приміщення має бути не менше, ніж 3.2м, а об'єм і площа - 15 м³ і 4.5 м² відповідно на кожного працівника [12]

Тривалість роботи за ПК визначається видом і категорією трудової діяльності. Згідно з положеннями розділу 5 [11] мають передбачатися:

перерви для відпочинку і вживання їжі (обідні перерви);

перерви для відпочинку і особистих потреб (згідно з трудовими нормами);

додаткові перерви, що вводяться для окремих професій з урахуванням особливостей трудової діяльності.

При 8-годинній денній робочій зміні залежно від характеру праці для операторів із застосуванням ЕОМ слід призначати регламентовані перерви для відпочинку тривалістю 15 хвилин через кожні дві години

Приміщення, в якому розроблявся програмний продукт розташоване на сьомому поверсі 21-поверхового будинку. Приміщення має однобічне природне освітлення та загальне штучне освітлення. Вікно орієнтовано на схід, площа засклення 30%. Стіни і стеля обклеєні світлими шпалерами, підлога вкрита середнім за яскравістю килимовим покриттям.

Приміщення загальною площею 24 м^2 , ширина якого складає 4 м , довжина – 6 м , висота стелі – $2,85\text{ м}$.

Виходячи з того, що в приміщенні працює одна людина, отримаємо наступні дані, наведені в таблиці 4.1[11]:

Таблиця 5.1. - Фактичні та нормативні значення площі та об'єму приміщення

Параметр приміщення	Нормативний	Фактичний
Площа, м^2	6 і більше	17.0
Об'єм, м^3	20 і більше	52.7

Виходячи з даних наведених в таблиці 5.1, зробимо висновок, що розміри приміщення задовольняють існуючим вимогам.

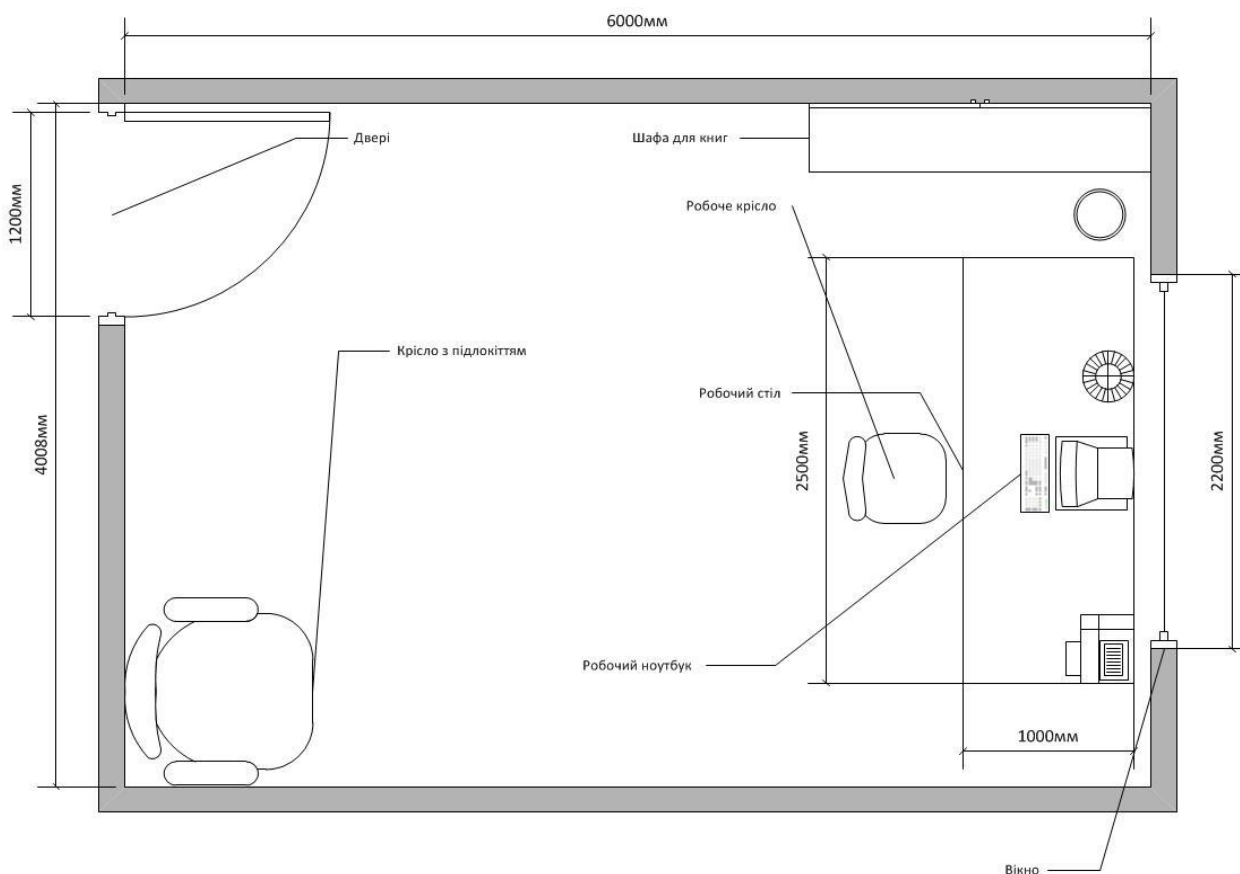


Рисунок 5.1 – План приміщення

Робоче місце знаходиться на відстані не менш, ніж на метр від стін з вікном. Габарити робочого місця оператора ЕОМ: висота столу – 800 мм, ширина – 1400 мм, глибина – 700 мм. Робочий стіл досить великий, але його висота не регулюються. Простір для ніг складає: висота – 750 мм, ширина – 600 мм, глибина на рівні колін – 550 мм, а на рівні простягнутої ноги – 750 мм. Робочий стілець оператора підйомно-поворотного типу, з регулятором висоти та повороту спинки. Висота спинки та її ширина мають розміри 500 мм та 380 мм відповідно. Висота спинки регулюються в межах 300 ± 20 мм, кут нахилу сидіння - до 15 град вперед і до 10 град назад. Кут нахилу спинки регулюються в межах 1-35 град від вертикального положення. Стілець також обладнаний стаціонарними підлокітниками довжиною 270 мм, та шириною 50 мм. Поверхня сидіння та спинки напівм'яка.

Екран монітора ноутбука розташований на відстані 700 мм від очей оператора. У зв'язку з тим, що ЕОМ, що використовуються в приміщенні –

ноутбуки, це не дозволяє змінювати кут нахилу клавіатури, але дозволяє змінювати кут нахилу монітору. Два робочі місця обладнані вмонтованими тумбами та шухлядами.

З обладнання в приміщенні розміщені:

- Ноутбук (AMD A6-3420M APU with Radeon(tm) HD Graphics (6x1.5 ГГц) / RAM 6 ГБ / HDD 500 ГБ / nVidia GeForce GT 710M, 2 ГБ / DVD Super Multi) з екранами 15.3" (1366x768, HD+ LED, глянцевої);
- Принтер HP LaserJet 3120 (A4, лазерний).

То ж характеристики робочого місця відповідають нормативним вимогам.

5.3 Мікроклімат

Мікроклімат приміщення - це сукупність параметрів повітря (температура, відносна вологість, повітрообмін, швидкість руху повітря, вміст у повітрі твердих часток (пилу), аромотерапія) у виробничому приміщенні, які діють на людину у процесі праці, на його робочому місці, у робочій зоні

Для постійних робочих місць, якими є робочі місця операторів ПК, встановлені оптимальні параметри мікроклімату, а при неможливості їх дотримання використовують допустимі параметри. Приміщення Іа категорії (виконуються легкі фізичні роботи), тому повинні дотримуватися наступні вимоги [11]:

Таблиця 5.2. – Параметри мікроклімату для приміщень з ПК

Період року	Параметр мікроклімату	Величина
Холодний	Температура повітря в приміщенні	22...24 °С
	Відносна вологість	40... 60%
	Швидкість руху повітря	0,1 м/с
Теплий	Температура повітря в приміщенні	23...25 °С
	Відносна вологість	40...60%
	Швидкість руху повітря	0,1 м/с

Для створення й автоматичної підтримки в приміщенні незалежно від зовнішніх умов оптимальних значень температури, вологості, чистоти і швидкості руху повітря, у холодний час року використовується водяне опалення, у теплий час року застосовується кондиціонування повітря. Кондиціонер являє собою вентиляційну установку, яка за допомогою приладів автоматичного регулювання підтримує в приміщенні задані параметри повітряного середовища.

Таблиця 5.3. – Норми подачі свіжого повітря в приміщення з ПК [13]

Характеристика приміщення	Об'ємна витрата свіжого повітря, що подається в приміщення, на одну людину в годину
Об'єм до 20 м ³ на людину	Не менше 30
20... 40 м ³ на людину	Не менше 20
Більше 40 м ³ на людину	Може бути використана природна вентиляція

5.4 Розрахунок системи кондиціонування

В кімнаті встановлено кондиціонер SAMSUNG AQ12ESG з охолоджувальною потужністю в 3,5кВт. Проведемо аналіз потужності за охолодження з урахуванням необхідної продуктивності за об'ємом повітря, що подається у приміщення, враховуючи всі теплові припливи і тепло виділення у приміщенні за умови найбільш несприятливих погодних умов.

Розрахунки продуктивності кондиціонера в робочому приміщенні проводиться за кількістю працюючих у приміщенні з урахуванням нормативних вимог [11]. У розрахунках враховується тепло, яке виділяється людьми і електроприладами. Вважається, що у спокійному стані людина виділяє 100 Вт

тепла, ПК – 300 Вт, інші прилади можна враховувати, що виділяють 1/3 своєї паспортної потужності.

Розраховується надлишок тепла. Для цього вираховуємо теплові потоки за формулою:

$$Q = Q_{\text{сонця}} + Q_{\text{ЕОМ}} + Q_{\text{людей}} = q * V_{\text{приміщення}} + 300 * n_{\text{ЕОМ}} + 100n_{\text{людей}} \quad (5.1)$$

Оскільки вікна виходять на південь, тому величина $q = 35$ становить Вт/м². Загальну величину теплового припливу знаходимо за формулою:

$$Q = 35 * 52,7 + 300 * 1 + 100 * 1 = 2244,5 \text{ Вт} \quad (5.2)$$

Кондиціонер повинен мати охолоджувальну потужність не менше значення, вказаного в розрахунку:

$$Q_{\text{сумарне}} = 1,2 * Q = 1,2 * 2244,5 = 2693,4 \text{ Вт} \quad (5.3)$$

Таким чином, кондиціонер, встановлений у приміщенні підходить для забезпечення належного охолодження в приміщенні.

5.5 Шум

Шум погіршує умови праці та негативно впливає на організм людини. Ті, хто працюють в умовах досить довгого шумового впливу, відчувають головні болі, запаморочення, зниження властивостей пам'яті, біль у вухах, тощо. Тривалий вплив інтенсивного шуму (більш ніж 80дБ) на слух людини може привести до часткової або навіть повної втрати слуху.

Приміщення знаходиться в будівля, що віддалена від гучних вулиць. Поблизу відсутні гучні об'єкти. Тому джерелом шуму може бути лише комп'ютерна техніка, що розташована в приміщенні, точніше її складові:

- Кулери блоків живлення;
- Кулери процесорів;
- Кулери відео карт;
- Жорсткий диск та дисководи.

Рівень звукового тиску і приміщені розраховується за формулою:

$$L_{\text{екв}} = 10 \lg \left(\frac{1}{T} * \sum t_i * 10^{0.1L_i} \right), \quad (5.4)$$

де T – загальний час дії системи, тут 8 годин,

t_i - час дії i -го елемента,

L_i - рівень звуку i -го елемента.

Основні джерела шуму та час впливу записано в таблиці 5.4.

Таблиця 5.4 – Джерела шуму

Джерело шуму	Рівень звуку, дБА	Час впливу, год	Кількість, шт.
Блок живлення	35	8	1
Кулер процесора	25	8	1
Кулервідеокарти	35	8	1
Жосткий диск	40	8	1
Дисковід	35	1	1
Зовнішній шум	45	8	1

Таким чином, еквівалентний рівень звукового тиску становить 48,17 дБА. Згідно [14] рівень звукового шуму не може перевищувати 50 дБА (для програмістів ЕОМ). Рівень шуму у приміщені задовольняє вимоги Держстандарту.

5.6 Випромінювання

Дисплей ноутбуку не є джерелом м'якого рентгенівського, інфрачервоного, ультрафіолетового, видимого, високо- і низькочастотного електромагнітного випромінювання. Джерелами ЕМВ є блоки рядкового та кадрового розгортання. Монітор задовольняє стандарту ТСО-99[15], що дозволяє підтримувати вище перераховані параметри в нормі.

5.7 Електробезпека

Джерелом можливо враження током для працівника на виробництві можуть слугувати: комп'ютер, принтер та інша офісна техніка, що живиться від електромережі.

ЕОМ є однофазовим споживачем електроенергії від трьохфазної мережі зі змінним током, напруженістю 380/220 В і частотою 50 Гц

В приміщенні розміщені такі споживачі електроенергії:

- Екран ноутбуку (1 шт.);
- Блок ноутбуку (1 шт.);
- Принтер (1 шт.);
- Світильник типу Philips(2 шт.).

Інформація з електробезпеки приміщення записані в таблиці 5.5.

Таблиця 5.5 – Дані з електробезпеки приміщення

№	Характеристика	Значення
1	Напруга електромережі	220 В
2	Електропроводка в приміщенні	Прихована (АПВ, АІ 5,25 мм)
3	Вологість	45-60%
4	Підлога	Ізолююча (ламінат)
5	Категорія	Без підвищеної

		електронебезпеки
--	--	------------------

Оскільки відносна вологість у приміщенні перебуває на рівні 50%, температура близько 24 градусів за Цельсієм, підлога – ламінат, непровідна, у приміщенні відсутні струмопровідний пил та хімічно активні речовини, то приміщення може бути віднесено до приміщень без підвищеної електронебезпеки.

У розглянутому приміщенні електропроводка схована, розетки встановлені на висоті 1м від підлоги. Висот, на якій встановлені освітлюючі лампи, складає 3 м.

Корпус, монітор і клавіатура ноутбука і принтер виготовлені з непровідного матеріалу, що робить неможливим враження електричним струмом при дотику до них. При виконанні робіт по ремонту і обслуговуванню ЕОМ обслуговуючий персонал зобов'язаний користуватися «правилами техніки безпеки при експлуатації електроустановок споживачами». При цьому при вході в приміщення має висіти плакат з інструкціями про правила безпеки при роботі з електрообладнанням. До роботи не допускаються особи, які не пройшли навчання по техніці безпеки.

5.8 Пожежна безпека

Конструкція будинку виконана із залізобетонних плит, тобто конструктивні елементи – неспалювані. Будинок відноситься до II ступеня вогнестійкості. Межа вогнестійкості конструкції 0,5-2,5 години. По вибухонебезпеки та пожежній небезпеці приміщення відноситься до категорії В, тому що в ньому знаходяться погано горючі, тверді та волокнисті речовини та матеріали: столи з пластмаси, ноутбуки, периферія, лінолеум, жалюзі. Згідно до правил розташування електроустановок, наше приміщення класифікується до II-ІІа класу, оскільки в приміщенні знаходяться тверді, спалювані речовини та матеріали [16].

Для підвищення безпеки, в тому числі пожежної безпеки, мережа напругою 220 В являється проводом з мідною жилою, що розташовується у вінілопластикових трубах, що прокладені в стінах. Ізоляція проводів розрахована на напругу в 1,5 кВ.

В приміщенні встановлена пожежна сигналізація Ю9078, сповіщувач Гром-12К, комбінований (з пульсуючою індикацією, що реагує на задимлення та температуру). Призначений для подачі світлових та звукових сигналів в системах охоронної та пожежної сигналізації. Сигналізація приміщення підключення до централізованого інформаційного пункту, від якого у випадку пожежі сигнал передається до пожежної станції. Крім цього, в приміщенні знаходяться один вогнегасник типу ВП-8 та два вогнегасники ВВК-1,4 [17].

Для приміщення, в якому працюють менше ніж 25 людей, та відстань до виходу евакуації від робочих місць не перевищує 25 метрів згідно норм достатньо одного евакуаційного виходу. Параметри евакуаційного виходу відповідають встановленим нормам : двері відкриваються назовні, ширина коридору 2 м, висота перекриття 3 м. Висота та ширина дверей у коридорі відповідно становлять 2,1 м та 2,2 м. Ширина дверей у приміщенні – 1,2 м.

Згідно вихідних даних і вимог чинних нормативних актів, з метою недопущення пожеж або зменшення небезпеки та шкоди, яку можуть нанести пожежі та вибухи, у приміщенні проводиться комплекс протипожежних заходів. Крім загальних вимог пожежної безпеки також здійснюють спеціальні заходи для окремих видів виробництв, технологічних процесів та промислових об'єктів:

- Не рідше 1 разу на квартал проводиться очищення агрегатів від пилу;
- Підготовки та перевірка шляхів евакуації. Кількість не менше двох. Двері мають відчинятися в напрямку виходу з будівлі. Встановлені протипожежні перешкоди.

5.9 Естетичні фактори виробничого середовища

Особливу роль в естетичній організації виробничого середовища відіграє колір, оскільки він справляє багатоплановий вплив на людину, а отже, має багатofункціональне призначення. Колір впливає на фізіологічні і психічні процеси, емоційні стани, працездатність і продуктивність праці працівників. Це зумовлюється такими характеристиками кольору, як колірний тон, насиченість (чистота) і яскравість (відображення світла).

Вибір колірної оформлення виробничих приміщень залежить від багатьох факторів — м'язових і нервових навантажень, температурного режиму, розмірів та орієнтації приміщення, монотонності роботи. Так, на роботах, які вимагають великих фізичних і нервових навантажень, а також у цехах з високою температурою повітря доцільно використовувати світлі тони голубого, зеленого та інших спокійних холодних кольорів невеликої насиченості. Якщо робота вимагає лише періодичних значних розумових і фізичних навантажень, то вона легше виконується у приміщеннях, пофарбованих у теплі кольори, які підвищують активність організму. Виконання монотонних робіт більш ефективно, якщо приміщення пофарбувати у яскраві кольори, які привертають увагу працівників і розширюють поле коркової активності. Світло-жовтий колір згідно з [18] урівнює емоційний стан.

5.10 Висновки до п'ятого розділу

На основі проведеного аналізу можна зробити висновки, що дане приміщення задовольняє вимогам [21]: площа на людину складає $17,00 \text{ м}^2$ та об'єм — $52,7 \text{ м}^3$, фактичний мікроклімат знаходяться в рамках норм. Система кондиціонування має достатню потужність, щоб зменшувати вплив додаткового тепла в приміщенні. Рівень шуму на рівні 48 дБА, що менше нормативного показника. В рамках пожежної безпеки встановлена сигналізація

та розміщено достатню кількість вогнегасників відповідних типів, характеристики шляхів евакуації відповідають нормам.

ВИСНОВКИ

У даній роботі було досліджено переміщення абітурієнтів при поданні документів до ВНЗ. Проаналізовано фактори, які впливають на вибір вступником навчального закладу, у якому він буде отримувати вищу освіту та наведено порядок визначення конкурентів ВНЗ для НТУУ "КПІ". У роботі сформовано вимоги до бази даних, у якій будуть зберігатись дані про вступників, та модулю, що завантажує дані з відкритого джерела інформації про абітурієнтів ІС "Конкурс". Розроблену систему було протестовано. Результати тестувань показали, що при додаванні даних про вступників до бази було використано правильні критерії, перевірка яких дозволяє запобігти колізій у БД.

Також систему було протестовано на швидкодію. Результати тестувань показали, що повне завантаження даних до БД відбувається протягом 1,5-2 годин, а при використанні більш потужного комп'ютера дані завантажуються близько 40 хвилин. Створену БД було проаналізовано на предмет повноти даних. Дані тестування свідчать про те, що дані у БД повністю відповідають тим, що доступні на ІС "Конкурс".

За допомогою розробленої системи було проведено чотири експерименти. Отримані дані у результаті проведення цих експериментів дозволили визначити ВНЗ-конкуренти для НТУУ "КПІ". Навчальні заклади, які можна назвати конкурентами НТУУ "КПІ" визначались у різних розрізах: за кількістю осіб, що подавали документи на усі можливі та окремо на технічні напрями, за кількістю заяв, та за сумою балів сертифікатів ЗНО і балу атестату про середню освіту. Також був проведений експеримент із застосуванням методів кластерного аналізу. За результатами цього експерименту можна зробити висновок, що вони підтвердили результати інших.

Отже, найбільшими конкурентами НТУУ "КПІ" можна назвати наступні навчальні заклади:

- НАУ;

- КНУ ім. Шевченка;
- НУ "ЛП".

Наведені заклади мають найбільшу кількість спільних вступників із НТУУ "КПІ". Експерименти показали, що більшість таких абітурієнтів віддає перевагу на користь НТУУ "КПІ". Причому кількість студентів першого курсу подекуди у п'ять разів вища, ніж до іншого ВНЗ.

У першому розділі роботи згадувалось, що ВНЗ сперечаються не тільки за кількість вступників, а й за тих, що мають високі конкурсні бали, адже для того, щоб надати високий рівень освіти для своїх студентів, вони мають бути добре підготовленими до навчання. Для того, щоб визначити перевагу ВНЗ серед сильних абітурієнтів, було проведено четвертий експеримент, який показав, що не дивлячись на більшу кількість вступників, відносна кількість сильних абітурієнтів віддає перевагу конкурентним закладам. Причому така тенденція спостерігається як серед загальної кількості вступників, так серед вступників на напрями ІТ-галузей.

Така ситуація може бути спричинена декількома факторами. По-перше, як відомо, кожен навчальний заклад має право самостійно визначати дисципліни, сертифікати ЗНО з яких вступник може подати на вступ до цього ВНЗ. Для всіх без винятку закладів обов'язковими екзаменами є українська мова та література і математика. Щодо третього екзамену, то на відміну від НТУУ "КПІ", який обов'язковим встановлює ЗНО з фізики, його конкуренти дають вступнику вибір: іноземна мова(будь-яка) або фізика. Якщо вступник має вищий бал з іноземної мови ніж з фізики, то задля збільшення конкурсного балу, ним буде подано сертифікат саме з іноземної мови. Цим можна пояснити те, що більшість вступників з найвищими балами вступають до інших ВНЗ.

По-друге, згідно з даними у створеній БД, НТУУ "КПІ" має у 1,5-2 рази більшу кількість бюджетних місць, ніж у його найбільших конкурентів. Це може означати те, що прохідний бал у таких ВНЗ може бути вищим, ніж у НТУУ "КПІ".

Отже, для НТУУ "КПІ" можна надати наступні рекомендації щодо заходів по залученню більшої кількості вступників:

1. Для тих технічних напрямів, де фізика не є фундаментальним предметом, надати абітурієнтам можливість вибору між фізикою та іноземною мовою. Це дасть змогу залучити більшу кількість сильних вступників на навчання за таким напрямом.

2. Для того, щоб залучити більшу кількість вступників на напрями, де фізика є обов'язковим предметом, а також які конкуренти теж мають, необхідно зробити більш гнучкими умови нарахування бонусних балів. Причому необхідно запропонувати такі умови, яких не пропонують конкурентні ВНЗ. Це дасть змогу тих простимулювати вступників, що здавали ЗНО з фізики та іноземної мови та мають бал з фізики нижчий.

3. Як відомо, НТУУ "КПІ" має курси доуніверситетської підготовки. Для того, щоб залучити більшу кількість абітурієнтів з регіонів України, необхідно організовувати курси підготовки у цих регіонах на базі шкіл із поглибленим вивченням математики та фізики. Це дасть змогу отримати перевагу над регіональними ВНЗ, як наприклад, НУ "ЛП".

4. Під патронатом НТУУ "КПІ" проводити різні конкурси серед випускників шкіл. Це дозволить звернути увагу вступників на НТУУ "КПІ".

Як видно з результатів роботи, вдалося визначити конкурентів КПІ. Але, раніше вказувалося, що аналіз проводиться тільки на даних вступних кампанії 2014 року. Якщо зробити аналіз даних попередніх років, то з'являється висока вірогідність того, що отримані результати будуть суттєво відрізнятися від наведених. Це пояснюється тим, що щороку умови вступу змінюються і стають дедалі складнішими. Також, кожен ВНЗ має можливість також встановлювати свої правила вступу і вони теж можуть суттєво впливати на вибір абітурієнта. Окрім факторів, зумовленими умовами вступу, є й інші фактори. Наприклад, виникнення різних надзвичайних ситуацій, що вимушують вступника обрати інший ВНЗ у іншому регіоні, або недостатня кількість коштів для життя у великому місті і навчання у більш престижному закладі. Виходячи з цього, не

можна стверджувати, що отримані результати враховують такі фактори та не є стовідсотково достовірними.

Тим не менше, створену базу даних та модуль, що завантажує дані до неї, можна використовувати кожному без винятку ВНЗ для визначення конкурентів. Також отримана БД також надає можливість отримати безліч іншої статистичної інформації, що дасть змогу будувати плани стратегічного розвитку навчального закладу. Це все можливо за умов, якщо розроблена система буде постійно підтримуватися та дороблюватися, оскільки умови вступу кожного року змінюються. Під зміни та модифікації також потрапляє і сайт ІС "Конкурс", звідки беруться дані. Якщо подивитись на те, як виглядав сайт, наприклад, у 2013 році та у 2014 році, то одразу ж можна буде помітити значні зміни. Нажаль, розроблена система не є стійкою до таких змін. Тому після завершення вступної кампанії наступного року, не можна буде отримати такі ж достовірні результати, як при роботі із даними 2014 року. Є два можливих вирішення цієї проблеми. Перший варіант - це необхідно мати групу розробників, що будуть вносити доробки до цієї системи задля постійного її функціонування. Іншим вирішенням може бути розробка більш універсального механізму завантаження даних, що самостійно знаходить зміни у структурі веб-сайту, самостійно адаптується під нові умови і змінює структуру бази даних.

ПЕРЕЛІК ПОСИЛАНЬ

1. Наказ "Про затвердження умов прийому до вищих навчальних закладів України у 2014 році" [Електронний ресурс]: наказ від 29 жовтня 2013 №1510 - Режим доступу: <http://zakon4.rada.gov.ua/laws/show/z1855-13> - Дата доступу: 28.04.2015.

2. Наказ "Про затвердження умов прийому до вищих навчальних закладів України у 2015 році" [Електронний ресурс]: наказ від 4 листопада 2014 №1172 - Режим доступу: <http://mon.gov.ua/content/%D0%94%D1%96%D1%8F%D0%BB%D1%8C%D0%BD%D1%96%D1%81%D1%82%D1%8C/%D0%92%D1%81%D1%82%D1%83%D0%BF%D0%BD%D0%B0%20%D0%BA%D0%B0%D0%BC%D0%BF%D0%B0%D0%BD%D1%96%D1%8F%202015/umovi-priyomu.rar> - Дата доступу: 28.04.2015.

3. Правила прийому до Національного технічного університету України «Київський політехнічний інститут» у 2014 році [Електронний ресурс]. - Режим доступу: <http://kpi.ua/rule> - Дата доступу: 28.04.2015.

4. Постанова "Про створення Єдиної державної електронної бази з питань освіти" [Електронний ресурс]: постанова від 13 липня 2011 №752-2011-п - Режим доступу: <http://zakon4.rada.gov.ua/laws/show/752-2011-%D0%BF> - Дата доступу: 28.04.2015.

5. Наказ МОНУ від "Про забезпечення функціонування інформаційної системи "Конкурс". [Електронний ресурс]: наказ від 11 липня 2008 №514 - Режим доступу: http://osvita.ua/legislation/Vishya_osvita/515/ - Дата доступу: 29.04.2015.

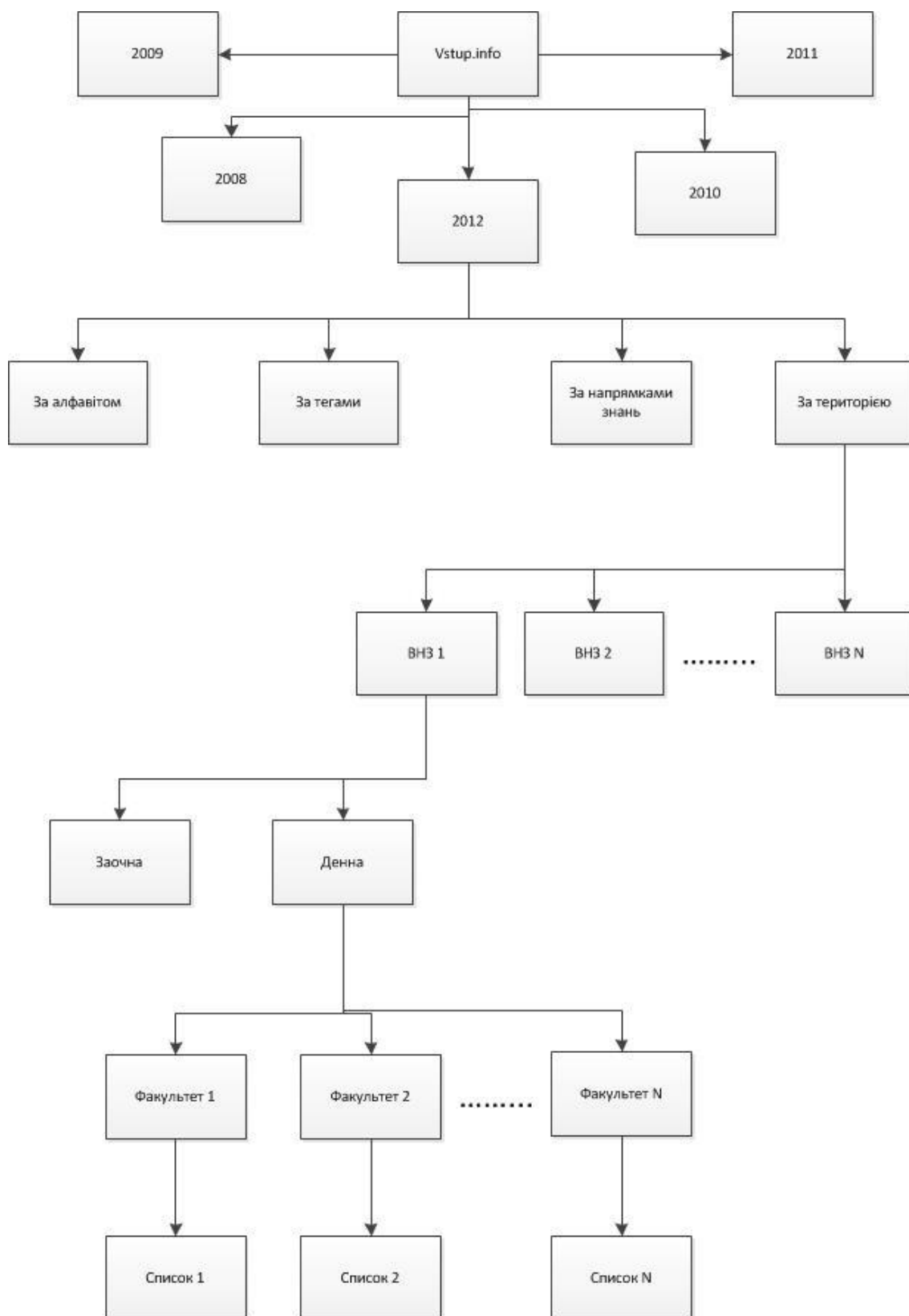
6. Закон України "Про захист персональних даних" [Електронний ресурс]: закон від 1 липня 2010 №2297-17 - Режим доступу: <http://zakon4.rada.gov.ua/laws/show/2297-17> - Дата доступу: 01.04.2013.

7. "Про проект "Компас"" [Електронний ресурс]: опис проекту "Компас" – Режим доступу: <http://bestuniversities.com.ua/ua/pro-proekt-0> - Дата доступу: 29.04.2015
8. "НТУУ «КПІ» знов у списку кращих університетів планети" [Електронний ресурс] текст - режим доступу: <http://kpi.ua/qs-15> - Дата доступу: 29.04.2015
9. "Методика визначення рейтингів "Топ-200 Україна"" [Електронний ресурс] текст - режим доступу: <http://www.euroosvita.net/index.php/?category=33> - дата доступу: 29.04.2015.
10. "Mathematica" [Електронний ресурс] текст - режим доступу: <https://ru.wikipedia.org/wiki/Mathematica> - дата доступу 12.05.2015.
11. Державні санітарні правила і норми роботи з візуальними дисплейними терміналами електронно-обчислювальних машин ДСанПіН 3.3.2.007-98 (затверджено Постановою Головного державного санітарного лікаря України від 10.12.1998 р. № 7).
12. Правила охорони праці під час експлуатації електронно-обчислювальних машин. НПАОП 0.00-1.28-10 (затверджено наказом Державного комітету України з промислової безпеки, охорони праці та гірничого нагляду від 26.03.2010р. № 65).
13. Санітарні норми мікроклімату виробничих приміщень. ДСН 3.3.6.042-99 (затверджено Постанова Головного Державного санітарного лікаря України від 1.12.1999 р. № 42)
14. Санітарні норми виробничого шуму, ультразвуку та інфразвуку [Електронний ресурс] : текст – Режим доступу:<http://dnop.com.ua/dnaop/act4878.htm>
15. Стандарт ТСО-99 [Електронний ресурс]: текст – Режим доступу: <http://school.iatp.by/roditel/kompivred/tso.htm>
16. Норми визначення категорій приміщень, будинків та зовнішніх установок за вибухопожежною та пожежною безпекою [Електронний ресурс] : текст – Режим доступу: <http://3umf.com/doc/5516/>

17. Про затвердження типових норм належності вогнегасників
[Електронний ресурс] : текст – Режим доступу:
<http://zakon4.rada.gov.ua/laws/show/z0554-04>

18. Крушельницька Я.В. Фізіологія і психологія: Підручник / Я.В.
Крушельницька //К.: КНЕУ, 2003. — 367 с.

ДОДАТОК А

Структура сайту <http://vstup.info>

ДОДАТОК Б

Процедури, що створюють структуру БД

1. Enrollments

```

CREATE TABLE [Enrollments] (
  [Id]          INT IDENTITY NOT NULL,
  [InstitutionId]  INT NOT NULL,
  [EnrollmentDirectionId] INT NOT NULL,
  [EnrolleeId]    INT NOT NULL,

  [RatingPosition]      SMALLINT NOT NULL,
  [TotalGrade]          FLOAT(24) NOT NULL,
  [CertificateGrade]    FLOAT(24) NOT NULL,
  [ChitGrade]          FLOAT(24) NOT NULL,

  /* TODO: Subjects (table) grades*/

  [Status]              NVARCHAR(20) NULL,
  [CoursesGrade]        FLOAT(24) NULL,
  [ExamsGrade]          FLOAT(24) NULL,
  [OlimpicGrade]        FLOAT(24) NULL,
  [OutOfContest]        BIT NOT NULL,
  [TopPriority]          BIT NOT NULL,
  [TargetDir]           BIT NOT NULL,

  [Originals]           BIT NOT NULL,
  [OriginalsToOther]    BIT NOT NULL,

  CONSTRAINT [PK_Enrollments] PRIMARY KEY CLUSTERED ([Id] ASC),
  CONSTRAINT [FK_EnrolleeEnrollment] FOREIGN KEY ([EnrolleeId]) REFERENCES [Enrollees] ([Id]),
  CONSTRAINT [FK_EnrollmentEnrollmentDirection] FOREIGN KEY ([EnrollmentDirectionId]) REFERENCES
[EnrollmentDirections] ([Id]),
  CONSTRAINT [FK_EnrollmentInstitution] FOREIGN KEY ([InstitutionId]) REFERENCES [Institutions] ([Id])
);

GO
CREATE NONCLUSTERED INDEX [IX_FK_EnrollmentInstitution]
  ON [Enrollments]([InstitutionId] ASC);

GO
CREATE NONCLUSTERED INDEX [IX_FK_EnrollmentEnrollmentDirection]
  ON [Enrollments]([EnrollmentDirectionId] ASC);

GO
CREATE NONCLUSTERED INDEX [IX_FK_EnrolleeEnrollment]
  ON [Enrollments]([EnrolleeId] ASC);

```

2. Enrollees

```

CREATE TABLE [Enrollees] (
  [Id]          INT NOT NULL IDENTITY,
  [Name]        NVARCHAR(30) NOT NULL,
  [Surname]     NVARCHAR(30) NOT NULL,
  [LastName]    NVARCHAR(30) NOT NULL,

```

```

CONSTRAINT [PK_EnrolleeSet] PRIMARY KEY CLUSTERED ([Id] ASC),
CONSTRAINT [UNIQUE_FIO] UNIQUE ([Name], [Surname], [LastName]),
);

```

3. EnrolleesDirections

```

CREATE TABLE [EnrollmentDirections] (
  [Id]          INT          IDENTITY NOT NULL,
  [TotalPlaces] SMALLINT NOT NULL,
  [BudgetPlaces] SMALLINT NOT NULL,
  [ApplicationCount] SMALLINT NOT NULL,
  [OriginalCount] SMALLINT NOT NULL,
  [DegreeId]    INT          NOT NULL,
  [DirectionId] INT          NULL,
  [DepartmentId] INT        NULL,
  [BranchId]    INT          NULL,
  [SpecialtyId] INT          NULL,
  [SpecializationId] INT    NULL,
  CONSTRAINT [PK_EnrollmentDirectionSet] PRIMARY KEY CLUSTERED ([Id] ASC),
  CONSTRAINT [FK_BranchEnrollmentDirection] FOREIGN KEY ([BranchId]) REFERENCES [Branches] ([Id]),
  CONSTRAINT [FK_DegreeEnrollmentDirection] FOREIGN KEY ([DegreeId]) REFERENCES [Degrees] ([Id]),
  CONSTRAINT [FK_DepartmentEnrollmentDirection] FOREIGN KEY ([DepartmentId]) REFERENCES [Departments]
  ([Id]),
  CONSTRAINT [FK_DirectionEnrollmentDirection] FOREIGN KEY ([DirectionId]) REFERENCES [Directions] ([Id]),
  CONSTRAINT [FK_SpecializationEnrollmentDirection] FOREIGN KEY ([SpecializationId]) REFERENCES
  [Specializations] ([Id]),
  CONSTRAINT [FK_SpecialtyEnrollmentDirection] FOREIGN KEY ([SpecialtyId]) REFERENCES [Specialties] ([Id]),
  CONSTRAINT [UNIQUE_EnrollmentDirection] UNIQUE ([TotalPlaces], [BudgetPlaces], [ApplicationCount],
  [OriginalCount], [DegreeId], [DirectionId], [DepartmentId], [BranchId], [SpecialtyId], [SpecializationId]),
);

```

```

GO
CREATE NONCLUSTERED INDEX [IX_FK_DegreeEnrollmentDirection]
  ON [EnrollmentDirections]([DegreeId] ASC);

```

```

GO
CREATE NONCLUSTERED INDEX [IX_FK_DirectionEnrollmentDirection]
  ON [EnrollmentDirections]([DirectionId] ASC);

```

```

GO
CREATE NONCLUSTERED INDEX [IX_FK_DepartmentEnrollmentDirection]
  ON [EnrollmentDirections]([DepartmentId] ASC);

```

```

GO
CREATE NONCLUSTERED INDEX [IX_FK_BranchEnrollmentDirection]
  ON [EnrollmentDirections]([BranchId] ASC);

```

```

GO
CREATE NONCLUSTERED INDEX [IX_FK_SpecialtyEnrollmentDirection]
  ON [EnrollmentDirections]([SpecialtyId] ASC);

```

```

GO
CREATE NONCLUSTERED INDEX [IX_FK_SpecializationEnrollmentDirection]
  ON [EnrollmentDirections]([SpecializationId] ASC);

```

4. Institutions

```

CREATE TABLE [Institutions] (
  [Id]          INT NOT NULL IDENTITY,
  [Title]       NVARCHAR(240) NOT NULL UNIQUE,
  CONSTRAINT [PK_Institutions] PRIMARY KEY CLUSTERED ([Id] ASC),
);

```

5. Specializations

```
CREATE TABLE [Specializations] (  
  [Id] INT IDENTITY NOT NULL,  
  [Name] NVARCHAR (40) NOT NULL UNIQUE,  
  CONSTRAINT [PK_Specializations] PRIMARY KEY CLUSTERED ([Id] ASC)  
);
```

6. Specialities

```
CREATE TABLE [Specialties] (  
  [Id] INT IDENTITY NOT NULL,  
  [Name] NVARCHAR (40) NOT NULL UNIQUE,  
  CONSTRAINT [PK_Specialties] PRIMARY KEY CLUSTERED ([Id] ASC)  
)
```

7. Directions

```
CREATE TABLE [Directions] (  
  [Id] INT IDENTITY NOT NULL,  
  [Name] NVARCHAR (40) NOT NULL UNIQUE,  
  CONSTRAINT [PK_Directions] PRIMARY KEY CLUSTERED ([Id] ASC)  
);
```

8. Departments

```
CREATE TABLE [Departments] (  
  [Id] INT IDENTITY NOT NULL,  
  [Title] NVARCHAR (240) NOT NULL UNIQUE,  
  CONSTRAINT [PK_Departments] PRIMARY KEY CLUSTERED ([Id] ASC)  
);
```

9. Degrees

```
CREATE TABLE [Degrees] (  
  [Id] INT NOT NULL IDENTITY,  
  [Name] NVARCHAR (60) NOT NULL UNIQUE,  
  CONSTRAINT [PK_Degrees] PRIMARY KEY CLUSTERED ([Id] ASC)  
);
```

10. Branches

```
CREATE TABLE [Branches] (  
  [Id] INT IDENTITY NOT NULL,  
  [Name] NVARCHAR (40) NOT NULL UNIQUE,  
  CONSTRAINT [PK_Branches] PRIMARY KEY CLUSTERED ([Id] ASC)  
);
```


ДОДАТОК В

Приклади запитів для формування даних для експериментів

1) Експеримент 1: Загальна кількість абітурієнтів

```

USE AdmissionDB;
select i.Title, count(distinct en.Id) as totalEnrollees

from Enrollments as et
inner join Enrollees as en on et.EnrolleeId = en.Id
inner join Institutions as i on et.InstitutionId = i.Id
inner join EnrollmentDirections as ed on et.EnrollmentDirectionId = ed.Id
inner join Degrees as dg on ed.DegreeId = dg.Id

where DegreeId = 3 and i.Title in( N'Національний технічний університет України "Київський політехнічний інститут",
N'Київський національний торговельно-економічний університет')
and et.OutOfContest = 0
group by i.Title
UNION
select N'Количество пересечений' as Collisions, count(*) as CollisionNumber

from (
    select distinct en.Id, count(distinct i.Title) as c1

    from Enrollments et
    inner join Enrollees en on et.EnrolleeId = en.Id
    inner join Institutions i on et.InstitutionId = i.Id
    inner join EnrollmentDirections as ed on et.EnrollmentDirectionId = ed.Id
    inner join Degrees as dg on ed.DegreeId = dg.Id

    where i.Title in( N'Національний технічний університет України "Київський політехнічний інститут",
N'Київський національний торговельно-економічний університет') and dg.id = 3
    and et.OutOfContest = 0
    group by en.Id
    ) as t1

where c1 > 1
UNION
select Concat(N'% от общего количества для ',i.Title),
10000*(select count(*) as CollisionNumber

from (
    select distinct en.Id, count(distinct i.Title) as c1

    from Enrollments et
    inner join Enrollees en on et.EnrolleeId = en.Id
    inner join Institutions i on et.InstitutionId = i.Id
    inner join EnrollmentDirections as ed on et.EnrollmentDirectionId = ed.Id
    inner join Degrees as dg on ed.DegreeId = dg.Id

    where i.Title in( N'Національний технічний університет України "Київський політехнічний
інститут"', N'Київський національний торговельно-економічний університет') and dg.id = 3
    and et.OutOfContest = 0
    group by en.Id
    ) as t1

```

```

where c1 > 1)/count(distinct en.Id) as Percentage

from Enrollments as et
inner join Enrollees as en on et.EnrolleeId = en.Id
inner join Institutions as i on et.InstitutionId = i.Id
inner join EnrollmentDirections as ed on et.EnrollmentDirectionId = ed.Id
inner join Degrees as dg on ed.DegreeId = dg.Id

where DegreeId = 3 and i.Title in( N'Національний технічний університет України "Київський політехнічний інститут",
N'Київський національний торговельно-економічний університет')
and et.OutOfContest = 0
group by i.Title
UNION
select CONCAT(N'Поступивших в ',i.Title), count(*) as NumberOfStudents

from Enrollments et
inner join Institutions i on i.Id = et.InstitutionId

where i.Title in( N'Національний технічний університет України "Київський політехнічний інститут", N'Київський
національний торговельно-економічний університет') and
et.EnrolleeId in(select u
from (
select distinct en.Id as u, count(distinct i.Title) as c1
from Enrollments et
inner join Enrollees en on et.EnrolleeId = en.Id
inner join Institutions i on et.InstitutionId = i.Id
inner join EnrollmentDirections as ed on
et.EnrollmentDirectionId = ed.Id
inner join Degrees as dg on ed.DegreeId = dg.Id
where dg.id = 3 and et.OutOfContest = 0 and i.Title in(
N'Національний технічний університет України "Київський політехнічний інститут", N'Київський національний
торговельно-економічний університет')
group by en.Id
) as t1
where c1 > 1)
and Originals = 1

group by i.Title

order by totalEnrollees desc

```

2) Експеримент 1: Загальна кількість абітурієнтів

```

USE AdmissionDB;
select i.Title, count(distinct en.Id) as totalEnrollees

from Enrollments as et
inner join Enrollees as en on et.EnrolleeId = en.Id
inner join Institutions as i on et.InstitutionId = i.Id
inner join EnrollmentDirections as ed on et.EnrollmentDirectionId = ed.Id
inner join Degrees as dg on ed.DegreeId = dg.Id
inner join Directions as d on d.Id = ed.DirectionId
where DegreeId = 3 and i.Title in( N'Національний технічний університет України "Київський політехнічний інститут",
N'Київський національний університет будівництва і архітектури')
and et.OutOfContest = 0 and d.Name in (N'системна інженерія',N'комп'ютерні науки',N'системний аналіз', N'програмна
інженерія', N'інформатика')
group by i.Title
UNION
select N'Количество пересечений' as Collisions, count(*) as CollisionNumber

```

```

from (
    select distinct en.Id, count(distinct i.Title) as c1

    from Enrollments et
    inner join Enrollees en on et.EnrolleeId = en.Id
    inner join Institutions i on et.InstitutionId = i.Id
    inner join EnrollmentDirections as ed on et.EnrollmentDirectionId = ed.Id
    inner join Degrees as dg on ed.DegreeId = dg.Id
    inner join Directions as d on d.Id = ed.DirectionId
    where i.Title in( N'Національний технічний університет України "Київський політехнічний інститут"',
N'Київський національний університет будівництва і архітектури') and dg.id = 3
    and et.OutOfContest = 0 and d.Name in (N'системна інженерія',N'комп'ютерні науки',N'системний
аналіз',N'програмна інженерія', N'інформатика')
    group by en.Id
    ) as t1

where c1 > 1
UNION
select Concat(N'% от общего количества для ',i.Title),
10000*(select count(*) as CollisionNumber

    from (
        select distinct en.Id, count(distinct i.Title) as c1

        from Enrollments et
        inner join Enrollees en on et.EnrolleeId = en.Id
        inner join Institutions i on et.InstitutionId = i.Id
        inner join EnrollmentDirections as ed on et.EnrollmentDirectionId = ed.Id
        inner join Degrees as dg on ed.DegreeId = dg.Id
        inner join Directions as d on d.Id = ed.DirectionId
        where i.Title in( N'Національний технічний університет України "Київський політехнічний
інститут"', N'Київський національний університет будівництва і архітектури') and dg.id = 3
        and et.OutOfContest = 0 and d.Name in (N'системна інженерія',N'комп'ютерні
науки',N'системний аналіз',N'програмна інженерія', N'інформатика')
        group by en.Id
        ) as t1

        where c1 > 1)/count(distinct en.Id) as Percentage

from Enrollments as et
inner join Enrollees as en on et.EnrolleeId = en.Id
inner join Institutions as i on et.InstitutionId = i.Id
inner join EnrollmentDirections as ed on et.EnrollmentDirectionId = ed.Id
inner join Degrees as dg on ed.DegreeId = dg.Id
inner join Directions as d on d.Id = ed.DirectionId
where DegreeId = 3 and i.Title in( N'Національний технічний університет України "Київський політехнічний інститут"',
N'Київський національний університет будівництва і архітектури')
and et.OutOfContest = 0 and d.Name in (N'системна інженерія',N'комп'ютерні науки',N'системний аналіз',N'програмна
інженерія', N'інформатика')
group by i.Title
UNION
select CONCAT(N'Поступивших в ',i.Title), count(*) as NumberOfStudents

from Enrollments et
inner join Institutions i on i.Id = et.InstitutionId

where i.Title in( N'Національний технічний університет України "Київський політехнічний інститут"', N'Київський
національний університет будівництва і архітектури') and
    et.EnrolleeId in(select u
        from (
            select distinct en.Id as u, count(distinct i.Title) as c1

            from Enrollments et
            inner join Enrollees en on et.EnrolleeId = en.Id
            inner join Institutions i on et.InstitutionId = i.Id

```

```

et.EnrollmentDirectionId = ed.Id
inner join EnrollmentDirections as ed on
inner join Degrees as dg on ed.DegreeId = dg.Id
inner join Directions as d on d.Id = ed.DirectionId
where dg.id = 3 and et.OutOfContest = 0 and d.Name in
(N'системна інженерія',N'комп'ютерні науки',N'системний аналіз',N'програмна інженерія', N'інформатика') and i.Title
in( N'Національний технічний університет України "Київський політехнічний інститут"', N'Київський національний
університет будівництва і архітектури')
group by en.Id
) as t1
where c1 > 1)
and Originals = 1
group by i.Title
order by totalEnrollees desc

```

3) Експеримент 3: Кількість спільних спеціальностей

Use AdmissionDB;

```
select N'Количество общих специальностей' as Collisions, count(*) as CollisionNumber
```

```
from (
```

```
select distinct d.Id, count(distinct i.Title) as c1
```

```
from Enrollments et
```

```
inner join Institutions i on et.InstitutionId = i.Id
```

```
inner join EnrollmentDirections as ed on et.EnrollmentDirectionId = ed.Id
```

```
inner join Degrees as dg on ed.DegreeId = dg.Id
```

```
inner join Directions as d on d.Id = ed.DirectionId
```

```
inner join Branches as b on b.Id = ed.BranchId
```

```
where i.Title in( N'Національний технічний університет України "Київський політехнічний інститут"',
N'Національний університет "Львівська політехніка") and dg.id = 3
```

```
and et.OutOfContest = 0 and BranchId in (16, 18, 14, 27, 13)
```

```
group by d.Id
```

```
) as t1
```

```
where c1 > 1
```

```
UNION
```

```
select Concat(N'% от общего количества для ',i.Title),
```

```
10000*(select count(*) as CollisionNumber
```

```
from (
```

```
select distinct d.Id, count(distinct i.Title) as c1
```

```
from Enrollments et
```

```
inner join Institutions i on et.InstitutionId = i.Id
```

```
inner join EnrollmentDirections as ed on et.EnrollmentDirectionId = ed.Id
```

```
inner join Degrees as dg on ed.DegreeId = dg.Id
```

```
inner join Directions as d on d.Id = ed.DirectionId
```

```
inner join Branches as b on b.Id = ed.BranchId
```

```
where i.Title in( N'Національний технічний університет України "Київський політехнічний
інститут"', N'Національний університет "Львівська політехніка") and dg.id = 3
```

```
and et.OutOfContest = 0 and BranchId in (16, 18, 14, 27, 13)
```

```
group by d.Id
```

```
) as t1
```

```
where c1 > 1)/count(distinct d.Id) as Percentage
```

```
from Enrollments as et
```

```
inner join Institutions as i on et.InstitutionId = i.Id
```

```
inner join EnrollmentDirections as ed on et.EnrollmentDirectionId = ed.Id
```

```

inner join Degrees as dg on ed.DegreeId = dg.Id
inner join Directions as d on d.Id = ed.DirectionId
inner join Branches as b on b.Id = ed.BranchId
where DegreeId = 3 and BranchId in (16, 18, 14, 27, 13) and i.Title in( N'Національний технічний університет України
"Київський політехнічний інститут", N'Національний університет "Львівська політехніка")
and et.OutOfContest = 0
group by i.Title

```

4) Експеримент 4: Визначення кількості вступників по сумі балів сертифікатів ЗНО та балу атестата про середню освіту.

```

Use AdmissionDB
select CONCAT(N'Поступивших среди больше 850 в ',i.Title), count(*) as NumberOfStudents

from Enrollments et
inner join Institutions i on i.Id = et.InstitutionId

where i.Title in( N'Національний технічний університет України "Київський політехнічний
інститут"', N'Національний університет "Львівська політехніка"') and
      et.EnrolleeId in(select u
                        from (
                            select distinct en.Id as u, count(distinct
i.Title) as c1
                            from Enrollments et
                            inner join Enrollees en on et.EnrolleeId =
en.Id
                            inner join Institutions i on
et.InstitutionId = i.Id
                            inner join EnrollmentDirections as ed on
et.EnrollmentDirectionId = ed.Id
                            inner join Degrees as dg on ed.DegreeId =
dg.Id
                            inner join Directions as d on d.Id =
ed.DirectionId
                            where dg.id = 3 and et.OutOfContest = 0 and
BranchId in (16, 18, 14, 27, 13) and i.Title in( N'Національний технічний університет
України "Київський політехнічний інститут"', N'Національний університет "Львівська
політехніка"')
                            group by en.Id
                            ) as t1
                        where c1 > 1)
      and Originals = 1 and TotalGrade > 850.5
group by i.Title
UNION
select CONCAT(N'Поступивших среди 800-850 в ',i.Title), count(*) as NumberOfStudents

from Enrollments et
inner join Institutions i on i.Id = et.InstitutionId

where i.Title in( N'Національний технічний університет України "Київський політехнічний
інститут"', N'Національний університет "Львівська політехніка"') and
      et.EnrolleeId in(select u
                        from (
                            select distinct en.Id as u, count(distinct
i.Title) as c1
                            from Enrollments et

```

```

en.Id inner join Enrollees en on et.EnrolleeId =
et.InstitutionId = i.Id inner join Institutions i on
et.EnrollmentDirectionId = ed.Id inner join EnrollmentDirections as ed on
dg.Id inner join Degrees as dg on ed.DegreeId =
ed.DirectionId inner join Directions as d on d.Id =
BranchId in (16, 18, 14, 27, 13) and i.Title in( N'Національний технічний університет
України "Київський політехнічний інститут"', N'Національний університет "Львівська
політехніка"')
group by en.Id
) as t1

where c1 > 1)
and Originals = 1 and TotalGrade between 849.5 and 800
group by i.Title
UNION
select CONCAT(N'Поступивших среди 750-800 в ',i.Title), count(*) as
NumberOfStudents

from Enrollments et
inner join Institutions i on i.Id = et.InstitutionId

where i.Title in( N'Національний технічний університет України "Київський політехнічний
інститут"', N'Національний університет "Львівська політехніка"') and
et.EnrolleeId in(select u
from (
select distinct en.Id as u, count(distinct
i.Title) as c1

from Enrollments et
inner join Enrollees en on et.EnrolleeId =
inner join Institutions i on
inner join EnrollmentDirections as ed on
inner join Degrees as dg on ed.DegreeId =
inner join Directions as d on d.Id =
where dg.id = 3 and et.OutOfContest = 0 and
BranchId in (16, 18, 14, 27, 13) and i.Title in( N'Національний технічний університет
України "Київський політехнічний інститут"', N'Національний університет "Львівська
політехніка"')
group by en.Id
) as t1

where c1 > 1)
and Originals = 1 and TotalGrade between 750 and 799.5
group by i.Title
UNION
select CONCAT(N'Поступивших среди 700-750 в ',i.Title), count(*) as
NumberOfStudents
from Enrollments et
inner join Institutions i on i.Id = et.InstitutionId

where i.Title in( N'Національний технічний університет України "Київський політехнічний
інститут"', N'Національний університет "Львівська політехніка"') and
et.EnrolleeId in(select u

```

```

from (
select distinct en.Id as u, count(distinct
i.Title) as c1
en.Id
et.InstitutionId = i.Id
et.EnrollmentDirectionId = ed.Id
dg.Id
ed.DirectionId
BranchId in (16, 18, 14, 27, 13) and i.Title in( N'Національний технічний університет
України "Київський політехнічний інститут"', N'Національний університет "Львівська
політехніка"')
group by en.Id
) as t1
where c1 > 1)
and Originals = 1 and TotalGrade between 700 and 749.5
group by i.Title
UNION
select CONCAT(N'Поступивших среди 650-700 в ',i.Title), count(*) as
NumberOfStudents
from Enrollments et
inner join Institutions i on i.Id = et.InstitutionId
where i.Title in( N'Національний технічний університет України "Київський політехнічний
інститут"', N'Національний університет "Львівська політехніка"') and
et.EnrolleeId in(select u
from (
select distinct en.Id as u, count(distinct
i.Title) as c1
en.Id
et.InstitutionId = i.Id
et.EnrollmentDirectionId = ed.Id
dg.Id
ed.DirectionId
BranchId in (16, 18, 14, 27, 13) and i.Title in( N'Національний технічний університет
України "Київський політехнічний інститут"', N'Національний університет "Львівська
політехніка"')
group by en.Id
) as t1
where c1 > 1)
and Originals = 1 and TotalGrade between 650 and 699.5
group by i.Title
UNION
select CONCAT(N'Поступивших среди 600-650 в ',i.Title), count(*) as
NumberOfStudents
from Enrollments et
inner join Institutions i on i.Id = et.InstitutionId

```

```

where i.Title in( N'Національний технічний університет України "Київський політехнічний
інститут"', N'Національний університет "Львівська політехніка"') and
      et.EnrolleeId in(select u
                        from (
i.Title) as c1
                        select distinct en.Id as u, count(distinct
                        from Enrollments et
en.Id
                        inner join Enrollees en on et.EnrolleeId =
                        inner join Institutions i on
et.InstitutionId = i.Id
                        inner join EnrollmentDirections as ed on
et.EnrollmentDirectionId = ed.Id
                        inner join Degrees as dg on ed.DegreeId =
dg.Id
                        inner join Directions as d on d.Id =
ed.DirectionId
                        where dg.id = 3 and et.OutOfContest = 0 and
BranchId in (16, 18, 14, 27, 13) and i.Title in( N'Національний технічний університет
України "Київський політехнічний інститут"', N'Національний університет "Львівська
політехніка"')
                        group by en.Id
                        ) as t1
                        where c1 > 1)
      and Originals = 1 and TotalGrade between 600 and 649.5
group by i.Title
UNION
      select CONCAT(N'Поступивших среди 550-600 в ',i.Title), count(*) as
NumberOfStudents
      from Enrollments et
inner join Institutions i on i.Id = et.InstitutionId

where i.Title in( N'Національний технічний університет України "Київський політехнічний
інститут"', N'Національний університет "Львівська політехніка"') and
      et.EnrolleeId in(select u
                        from (
i.Title) as c1
                        select distinct en.Id as u, count(distinct
                        from Enrollments et
en.Id
                        inner join Enrollees en on et.EnrolleeId =
                        inner join Institutions i on
et.InstitutionId = i.Id
                        inner join EnrollmentDirections as ed on
et.EnrollmentDirectionId = ed.Id
                        inner join Degrees as dg on ed.DegreeId =
dg.Id
                        inner join Directions as d on d.Id =
ed.DirectionId
                        where dg.id = 3 and et.OutOfContest = 0 and
BranchId in (16, 18, 14, 27, 13) and i.Title in( N'Національний технічний університет
України "Київський політехнічний інститут"', N'Національний університет "Львівська
політехніка"')
                        group by en.Id
                        ) as t1
                        where c1 > 1)
      and Originals = 1 and TotalGrade between 550 and 599.5
group by i.Title
UNION

```



```
        where c1 > 1)
    and Originals = 1 and TotalGrade < 499.5
group by i.Title
```