## МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ “КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ”**

Факультет прикладної математики

Кафедра програмного забезпечення комп’ютерних систем

## КУРСОВИЙ ПРОЕКТ ТЕХНІЧНЕ ЗАВДАННЯ

з дисципліни “Бази даних”

напряму підготовки 6.050103 – Програмна інженерія

на тему:

(назва теми)

|  |  |  |
| --- | --- | --- |
| **Студент**  **групи** КП-51 | (ПІБ) | (підпис) |
| **Викладач**  **к.т.н, доцент кафедри СПіСКС** | **Петрашенко А.В.** | (підпис) |

Київ–2018

1. **НАЙМЕНУВАННЯ ТА ГАЛУЗЬ ЗАСТОСУВАННЯ РОЗРОБКИ**

*Найменування*: Моніторингова система контролю успішності спортсменів

*Галузь застосування розробки та функціонал*: збір, фільтрація та аналіз результатів різних вікових категорій та рівня підготовки з метою прогнозування та вдосконалення спортивного процесу.

1. **ДАТА ПОЧАТКУ ТА ЗАКІНЧЕННЯ ПРОЕКТУ**

*Підготовка технічного завдання на затвердження*: 23 березня 2018 року.

*Захист курсового проекту*: 18 травня 2018 року.

1. **МЕТА РОЗРОБКИ**

Метою розробки даного курсового проекту є вдосконалення навичок роботи з обраною мовою програмування, відповідних бібліотек та СКБД; створення програмного застосунку, який дозволив би проводити аналіз даних спортивних змагань.

1. **ВИМОГИ ДО ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ**

Вимоги до структурних елементів системи:

*Генерація даних*: у системі має бути наявна можливість створення або завантаження даних про результати спортивних змагань.  
Для наповнення даних у СКБД для цієї системи буде використане надане спортивною організацією **© “World Archery”** **API (далі – WA API)**, за допомогою якого можна буде отримати результати проведених змагань. Завантажувані дані відображають результати реальних проведених спортивних змагань, отже, показуватимуть реальний стан речей.

*Фільтрація та валідація даних*: За допомогою наданого WA API можна буде проводити фільтрацію даних, відбираючи результати конкретних проведених змагань. Надані результати отримуються у форматі .json, і вони вже є валідованими. Окрім цього, надані результати містять велику кількість додаткових даних, які не є такими важливими для проектованої системи, і ці дані будуть відфільтровуватися (відкидуватися).

*Засоби реплікації та масштабування*: реплікація та масштабування будуть забезпечуватися засобами СКБД, яка буде використовуватися при розробці даного програмного застосунку (MongoDB).

MongoDB може працювати з набором реплік. Такий набір складається з двох або більше копій даних. Кожний з екземплярів набору реплік може в будь-який момент виступати у ролі основної або допоміжної репліки. Усі операції читання й запису за замовченням виконуються з основною реплікою. Допоміжні репліки підтримують в актуальному стані копії даних. У випадку, коли основна репліка дає збій, набір реплік вибирає, яка з реплік має стати основною. Вторинні репліки можуть додатково бути джерелом для операцій читання.

MongoDB масштабується горизонтально, використовуючи шардінг. Користувач обирає ключ шарда, який визначає, як будуть розподілені дані у колекції. Дані розподіляються на діапазони (в залежності від ключа шарда) та розподіляються по шардам.

*Аналіз даних*: за допомогою аналізу даних зможуть бути відстежені, наприклад, рекордні досягнення спортсменів або непомітні “неозброєним” оком залежності у результатах їх кар’єри. Для цього будуть використані відповідні алгоритми, представлені бібліотеками. Зокрема будуть написані власні алгоритми для проведення специфічного аналізу результатів.

*Резервування та відновлення даних*: дані, з якими відбувається робота у даному програмному застосунку, мають резервуватися. Таким чином, при виникненні проблем з основним джерелом даних, можна було б їх відновити з резервованого раніше набору даних. Для резервування та відновлення даних будуть використані засоби СКБД, яка буде використовуватися при розробці цього програмного застосунку (MongoDB). *Більш детально цей пункт буде описано пізніше.*

1. **ОБҐРУНТУВАННЯ ВИБОРУ СКБД**

Для даного програмного застосунку була вибрана СКБД MongoDB. MongoDB –документно-орієнтована СКБД, класифікована як NoSQL, яка не потребує опису схеми таблиць. Дана СКБД була вибрана завдяки своїм визначним можливостям, зокрема:

* *Спеціальні запити до БД*: Запити до бази даних можуть повертати конкретні поля документів або користувацькі JavaScript-функції. Також, існує підтримка пошуку з використанням регулярних виразів. Окрім цього, можна налаштувати запит, який повертатиме випадковий набір результатів (це може бути використано для отримання варіаційних наборів результатів для тестування або для відслідковування цікавих закономірностей у нестандартних наборах даних)
* *Реплікація та балансування навантаження*: див. п. 4, *Засоби реплікації та масштабування*.
* *Файлове сховище*: MongoDB може бути використана у якості файлового сховища з балансуванням навантаження та реплікацією даних. Ця функція, названа Grid File System, надається разом з драйверами MongoDB. MongoDB пропонує розробникам функції для роботи з файлами та їх вмістом. GrdFS розділяє файл на частини та зберігає кожну його частину як окремий документ.
* *Виконання JavaScript на стороні серверу*: JavaScript може бути використаний у запитах та відправлений БД для виконання.

1. **ВИМОГИ ДО ІНТЕРФЕЙСУ КОРИСТУВАЧА**

Об’єм розробки, необхідний для виконання даної курсової роботи у належному обсязі, не включає в себе створення належного інтерфейсу для користувача. Таким чином, дане програмне забезпечення розробляється не як продукт для кінцевого користувача, а як дослідницький проект, за допомогою передачі аргументів у який можна буде отримувати результат. Таким чином, мінімальний необхідний інтерфейс може бути забезпечений у консольному вигляді. При цьому, не виключається створення більш якісного інтерфейсу, який був би зрозумілим і приємнішим для використання будь-якими типами користувачів.

1. **ВИБІР ЗАСОБІВ РОЗРОБКИ**

Серед засобів розробки були виділені такі компоненти:

*Мова програмування*: Python 3.6. Зокрема, передбачається використання таких фрейморків та бібліотек:

* Бібліотеки напрямку Data Science:
  + OpenCV;
  + numpy;
  + scipy;
  + pandas;
  + та інші.
* Бібліотека для взаємодії з СКБД MongoDB: PyMongo;
* Інструментарій для Data Mining: Theano;
* Інструментарій для Big Data: Scikit-Learn;
* Можливе додаткове використання фреймворку *Django* для створення серверу з веб-інтерфейсом.

*СКБД*: MongoDB;

*IDE*: PyCharm Ultimate Edition;

*Джерело даних*: WA API.

1. **ЕТАПИ РОЗРОБКИ**

|  |  |  |
| --- | --- | --- |
| **Номер** | **Назва етапу розробки** | **Дата початку** |
| 1 | Початок роботи над курсовим проектом | 16.03.18 |
| 2 | Формування ТЗ | 23.03.18 |
| 3 | Проектування | 25.03.18 |
| 4 | Розроблення ПЗ | 08.04.18 |
| 5 | Розроблення документації | 14.05.18 |
| 6 | Тестування | 16.05.18 |
| 7 | Експлуатація та супровід | 18.05.18 |
| 8 | Захист курсового проекту | 18.05.18 |