

**Назва:** " Веб-кравлер для аналізу оголошень про роботу за допомогою веб-скрапінгу та Data Mining"

**Керівник (product owner):** Каук В.І.

**Проблема:**

- Великі обсяги даних про вакансії в інтернеті
- Складно відстежити зміни вимог роботодавців, рівня зарплат тощо
- Відсутність зручних аналітичних інструментів

**Анотація:**

- Веб-кравлер для автоматичного збору оголошень з сайтів
- Модуль веб-скрапінгу для парсингу даних
- Модуль аналітики та Data Mining для пошуку інсайтів
- Результати у вигляді звітів, дашбордів, аналітичних довідників

**Початкові дані:**

- Веб-сторінки з вакансіями (hh.ua, work.ua, dou.ua тощо)
- Структуровані дані компаній, IT-сектору

**Джерела аналогів:**

- ParseHub - веб-скрапінг та аналітика даних
- Octoparse - веб-скрапінг та парсинг даних

**Команда:**

- Продукт-менеджер
- Python-розробник - Scrapy, BeautifulSoup, Selenium
- Аналітик даних - Pandas, Matplotlib, NLP
- QA інженер
- Ресерчер даних

**Технології:**

- Python, Scrapy, BeautifulSoup, Selenium
- Jupyter Notebook для аналітики
- MySQL/Postgres для зберігання даних
- Docker для розгортання

**Критерії:**

- Обсяг даних > 1 млн вакансій
- Похибка аналізу < 5%

**Графік:**

- Листопад-Грудень 2023: Розробка парсера
- Січень-Лютий 2024: Data mining та аналітика
- Березень 2024: Тестування
- Квітень 2024: Запуск першої версії

Програмний продукт, який отримано протягом навчального процесу може бути використаний закладом вищої освіти за призначенням безоплатно і без обмежень у часі.