ПРОЕКТ № 239

[BUCKS](https://github.com/PTrendafilov/BUCKS)

Направление: Големи обеми от данни

**АВТОРИ:**

1. **Пламен Лъчезаров Трендафилов**
   1. ЕГН: 0641216485
   2. Адрес: София, ж.к Лагера, бл. 44, вх. А, ет. 5, ап. 16
   3. Телефон: 0876067114
   4. [p.trendafilova@abv.bg](mailto:p.trendafilova@abv.bg)
   5. НПМГ “Акад. Любомир Чакалов”
   6. 12И
2. **Иван Георгиев Иванов**
   1. ЕГН: 0648036382
   2. Адрес: София, ул. Маестро Кънев 61, ет.5, ап.22
   3. Телефон: 0884304615
   4. <ivanov.vang@gmail.com>
   5. Училище: НПМГ “Акад. Любомир Чакалов”
   6. Клас: 12И

**РЪКОВОДИТЕЛ:**

1. Екатерина Димитрова Мицева
   1. телефон: 0889389662
   2. ekaterina.mitseva@npmg.org
   3. длъжност: учител по информатика и ИТ

**BUCKS**

<https://github.com/PTrendafilov/BUCKS>

Проблем

Съществува сериозно предизвикателство сред инвеститорите и финансовите институции, когато става въпрос за анализиране на огромни обеми от пазарни и новинарски данни. От една страна, пазарът е динамичен и често реагира на политически и икономически новини. От друга страна, обобщаването на такава информация и вземането на решения в реално време изисква сериозна инфраструктура и автоматизация.

 Решение

Нашият прототип BUCKS е инвестиционна система, която събира, структурира и анализира новини и ценови данни от различни източници, като прилага модели за обработка на естествен език (NLP) и алгоритми с подсилване (Reinforcement Learning, RL). Това позволява бързо идентифициране на потенциални тенденции и генериране на препоръки с цел подобрена възвръщаемост.

Цели

1. Подобряване на скоростта и точността при анализ на финансови новини – чрез автоматично класифициране и оценка на влиянието на събитията.

2. Създаване на работещ прототип, който генерира препоръки – системата трябва да обобщава огромни обеми от данни и да предлага конкретни действия (купува, продава, задържа).

3. Изграждане на устойчива инфраструктура – за да може да се пренася и надгражда лесно в различни производствени среди.

 Целева група

- **Екипи за финансови анализи:** аналитични отдели в банки или брокерски къщи, които се нуждаят от бързо обобщение на пазара.

- **Независими инвеститори,** търсещи изкуствен интелект, за да оптимизират своите портфейли и да получават автоматични предложения.

Основни етапи в реализирането на проекта

1. **Проучване на съществуващи решения -** Извършване на анализ на вече налични системи за пазарен анализ и проучване на използваните NLP и RL алгоритми.

2. **Съставяне на план за реализация** - Избор на архитектура, бази данни и ключови технологии (Python, базови NLP библиотеки, Yahoo Finance API).

3. **Определяне на функциите, които системата ще изпълнява -** Приемане на новини от [Newsapi.ai](http://Newsapi.ai), интеграция с Yahoo Finance за ценови данни и свързването им към модул за машинно обучение.

4. **Планиране на структура и архитектура -** Оформяне на данните в база, дефиниране на отделните модули: Data Ingestion, NLP, RL и интерфейс за наблюдение.

5. **Проучване на необходимите технологии -** Анализ на възможностите за обработка на големи данни, автоматизация на заявки и трениране на модели.

6. **Създаване и разработване на основните части** - Писане на скриптове за извличане на новини и пазарни данни, прилагане на NLP слоеве и RL алгоритми за решения.

7. **Тестване и проверка за грешки -** Пробни бектестове и проверка на надеждността на NLP модула. Техническа стабилност на базата данни и API връзките.

8. **Публикуване в интернет -** Стартиране на прототипа на публичен сървър или облачна инфраструктура.

9. **Въвеждане в експлоатация** - Провеждане на тестове в реални условия (paper trading среди или ограничена реална търговия).

10. **Популяризиране** - Демонстриране на резултатите и представяне пред потенциални инвеститори или заинтересовани анализатори.

Ниво на сложност на проекта

Проектът е с високо ниво на сложност, поради необходимостта да се обхванат няколко технологични области едновременно – обработка на естествен език, потоци от пазарни данни и алгоритми с дълбоко обучение. Освен това темата за инвестиции предполага нужда от стабилна система за управление на риска и точност на данните. Необходими са както програмистки умения, така и разбиране на финансовата среда и статистическия анализ.

Съвместна работа с финансисти и експерти по данни

Поради специфичната сфера на проекта, ние търсим помощ и консултации с финансисти, икономисти и експерти в областта на машинното обучение. Например, дискусии с професионалисти, занимаващи се с търговия на стоки, петрол и зърнени култури. Разчитаме на ценни насоки за това кои индикатори и фактори действително оказват влияние. Също така работим със специалисти по големи данни, за да гарантираме мащабируемостта и ефективността на системата.

В приложените файлове се намира и документ, в който се споменава консултация с финансови експерти за валидиране на модели и тестови сценарии.

Логическо и функционално описание на решението

**1. Извличане на пазарни и новинарски данни**

Системата събира финансова информация (цени на петрол, пшеница и други активи) от Yahoo Finance и политико-икономически новини от [Newsapi.ai](http://Newsapi.ai).

**2. Обработване на език (NLP)**

   Новините се анализират, за да се определи дали са позитивни, неутрални или негативни спрямо даден пазар. Извличат се ключови теми, дати и контекст.

**3. Обединяване на данните**

   Резултатите от NLP се комбинират с текущите пазарни цени в база данни, като записите се обогатяват с индикатори (RSI, Moving Averages).

**4. Reinforcement Learning (RL) слой**

   Моделът взема предвид историческите зависимости между новини и ценови движения, за да генерира сигнали (купуване, продаване, задържане).

**5. Интерфейс и наблюдение**

   Активната част на системата подава сигнали за сделки в симулирана среда (paper trading) или реални брокерски платформи. Предвиждат се статистики и метрики за проследяване на печалби/загуби и рискови показатели.

Реализация

- **Клиентска част и потребителски интерфейс**:

  Използваме уеб приложение, изградено с HTML, CSS, JavaScript и Bootstrap шаблон. Това позволява удобен достъп до табла за наблюдение и тестови отчети.

- **Сървърна логика**:

  Основните модули за обработка са разработени с помощта на Python (Django) и Node.js за отделни микросървъри, където е необходимо.

- **Обработка и анализ на данни:**

  Натоварваме Python библиотеки като `spaCy` или `transformers` за NLP и `Stable Baselines3` (или подобни) за RL.

- **База данни**:

  SQLite е използвана в прототипния етап за по-лесно вграждане, но при нужда може да се премине към PostgreSQL за по-големи обеми.

- **Визуална идентификация**:

  Разработено е лого „BUCKS“ за разпознаваемост на проекта и последващи рекламни материали.

- **Анимирани демонстрации**:

  Част от рекламните клипове и презентации са създадени с инструменти като Animaker и качени в папката `video` или на слайдовете от презентацията.

За нуждите на демонстрация сме подготвили тестов `admin` панел в Django, който дава възможност за следене на записите в базата.

Развитие на приложението

В бъдеще планираме да публикуваме мобилна версия на BUCKS в Google Play и App Store. Така потребителите ще могат да получават известия за пазарни промени и новини директно на телефоните си.

Заключение

BUCKS има потенциала да улесни ежедневната работа на инвеститори и финансови анализатори, като позволява бързо и надеждно синтезиране на големи обеми от данни. Основната ни цел е да предоставим инструмент, който да помага за вземането на по-информирани решения, намаляване на риска и оптимизиране на търговските стратегии. С този прототип се надяваме да допринесем за развитието на автоматизираните инвестиционни системи и да покажем възможностите на съвременните техники за машинно обучение в реални пазарни условия.