

Универзитет „Свети Кирил и Методиј“ – Скопје
Факултет за информатички науки и компјутерско инженерство

Архитектура и имплементација на систем за обработка на финансиски податоци во реално време

Изработила:
Ангела Ангелеска, 221083

Содржина

Опис	3
Data Ingestion Service	3
Real-Time Processing Service	4
Data Storage Service	4
Data Analysis Service	5
Data Visualization Service	5
Функциски барања.....	6
Нефункционални барања.....	6
Use Case	7
Архитектура на апликацијата	8
1. Концептуална архитектура.....	8
2. Извршна архитектура	9
3. Имплементациона архитектура	10

Опис

Проектот се фокусира на создавање комплетен систем кој ќе обработува финансиски податоци од различни извори, трансформирајќи ги во корисни информации за анализа и визуелизација. Овој систем овозможува автоматизирано и ефикасно управување со податоци, почнувајќи од нивното прибирање, преку обработка и анализа, до нивното прикажување на крајни корисници.

Системот се состои од пет главни сервиси, секој со специфична улога во обработката, анализата и визуелизацијата на финансиски податоци:

1. Data Ingestion Service

Овој сервис собира финансиски податоци од различни надворешни извори преку автоматизирани процеси. Податоците се преземаат и се подготвени за понатамошна обработка и анализа.

2. Real-Time Processing Service

Овој сервис ги обработува податоците веднаш штом се добијат, користејќи технологии за стриминг како Apache Kafka и Apache Spark Streaming. Неговите задачи вклучуваат филтрирање, агрегација и трансформација на податоците, овозможувајќи понатамошно користење на истите.

3. Data Storage Service

Овој сервис е одговорен за складирање на обработените податоци, обезбедувајќи сигурност, интегритет и брз пристап до информациите. Тој овозможува зачувување на структурирани податоци кои се подготвени за анализа и визуелизација, како и лесно управување со историските и тековните информации.

4. Data Analysis Service

Овој сервис користи модели за машинско учење за анализа на податоци, овозможувајќи идентификација на трендови, предвидување идни сценарија и проценка на влијанието на различни фактори. Примената на овие напредни алгоритми обезбедува генерирање на точни и релевантни извештаи за поддршка при донесување одлуки.

5. Data Visualization Service

Овој сервис обезбедува веб-интерфејс кој ги претставува резултатите од анализата преку интерактивни графикони и табели. Визуелизацијата помага корисниците лесно да ги интерпретираат сложените финансиски информации.

Data Ingestion Service

Data Ingestion Service претставува клучна компонента на системот која обезбедува централизирано собирање и обработка на финансиски податоци и вести од надворешни извори, како што се Yahoo Finance и MarketAux API. Овој сервис креира REST API преку кој останатите сервиси лесно можат да пристапат до историските податоци и вестите за акциите.

Податоците за акции се собираат преку Yahoo Finance API кое овозможува повлекување на финансиски информации врз основа на даден код. Овие податоци опфаќаат временски период од десет години и се враќаат во формат погоден за

понатамошна обработка. Вестите за акции се комбинираат од повеќе извори, вклучувајќи ги Yahoo Finance и MarketAux API, со цел да се обезбеди комплетен преглед на релевантни информации. Собраните податоци се враќаат во стандарден формат кој лесно може да се користи од клиентски апликации или други сервиси.

API-то дефинира рути преку кои може да се пристапи до историските податоци и вестите. Овие рути овозможуваат обработка на барања и доставување на информации во стандарден формат, погоден за клиентски апликации или други сервиси во системот.

Real-Time Processing Service

Real-Time Processing Service е сервис кој обработува финансиски податоци и вести во реално време. Главната цел е да се обезбеди трансформација, филтрирање и организирање на податоците, кои потоа се испраќаат назад до Kafka за понатамошна употреба од други сервиси. Овој сервис користи Apache Kafka за пренос на податоци и Apache Spark Streaming за обработка на стримови.

Првично, сервисот собира кодови на компании од надворешни извори, кои се користат за идентификација на компаниите. Финансиските податоци и вестите за овие компании се преземаат преку API повици од претходно креираниот Data Ingestion Service, кој ги собира податоците од надворешни извори како Yahoo Finance и MarketAux API. За секоја компанија се генерираат пораки кои содржат историски податоци или релевантни вести. Овие пораки се серијализираат во JSON формат и се испраќаат до Kafka продуцерите. Потрошувачите во системот, имплементирани со Apache Spark Streaming, ги читаат пораките од Kafka topics и ги обработуваат во реално време. Обработените информации стануваат достапни за други компоненти во системот, како што се сервисите за анализа или визуелизација, така што информациите се праќаат до одредени потрошувачи кои се претплатени на излезните Kafka topics.

Data Storage Service

Data Storage Service е клучна компонента на системот која е одговорна за складирање на обработените финансиски податоци и вести. Овој сервис обезбедува сигурност, интегритет и брз пристап до информациите, овозможувајќи лесно управување со историските и тековните податоци. Складирањето е организирано така што податоците се зачувуваат во релациони бази на податоци, овозможувајќи нивна лесна употреба за анализа и визуелизација.

Data Storage Service создава посебни табели за секоја компанија во базата на податоци. Табелите ги содржат сите релевантни финансиски информации, како што се датум, open price, close price, high price, low price, volume, dividend и stock splits. Овие табели се креираат динамички, врз основа на кодовите добиени од претходните сервиси.

Покрај финансиските податоци, сервисот складира и вести поврзани со компаниите во посебни табели. Секоја табела содржи наслови и резимеа на вестите,

организирани според кодот на компанијата. Ова овозможува лесно пребарување и анализи на релевантните информации.

Обработените финансиски податоци и вести се примаат преку потрошувачи претплатени на одредени Kafka topics од Real-Time Processing Service. Потрошувачите во Data Storage Service ги читаат овие пораки и ги внесуваат во соодветните табели.

Data Analysis Service

Data Analysis Service е аналитички сервис во системот, дизајниран да обработува финансиски податоци и вести со цел да генерира сигнали за тргување, да идентификува трендови и да даде препораки како „Buy“, „Sell“ или „Hold“. Сервисот користи комбинација од технички индикатори, sentiment analysis и модели за машинско учење за да обезбеди точни и релевантни резултати. Овие аналитички податоци се интегрирани со останатите компоненти во системот, овозможувајќи поддршка при донесување одлуки и визуелизација.

Сервисот анализира историски финансиски податоци преку пресметка на технички индикатори како RSI, MACD, Bollinger Bands и други. Овие индикатори се користат за идентификување на трендови во пазарот и генерирање сигнали за тргување. Со помош на sentiment analysis, сервисот ги обработува вестите поврзани со компаниите, категоризирајќи ги како „Good news“, „Bad news“ или „Neutral news“. Врз основа на оваа анализа се генерираат препораки за тргување, кои се дополнително комбинирани со предвидувањата од моделите за машинско учење. Овие модели вршат предвидување на идните цени на акциите врз основа на историските податоци, анализирајќи карактеристики како open price, close price и volume.

Сервисот ги комбинира резултатите од сите направени предвидувања, ги обработува и генерира финален сигнал („Buy“, „Sell“ или „Hold“) кој е најсоодветен според трендовите во пазарот и sentiment analysis од вестите. Обработените податоци се лесно достапни преку API рути.

Data Visualization Service

Data Visualization Service е завршната компонента на системот, која обезбедува интерактивен и интуитивен веб-интерфејс за визуелизација на финансиските податоци и аналитички резултати. Овој сервис е дизајниран да ги прикаже обработените информации во форма на графикони и табели, овозможувајќи лесна интерпретација на сложените финансиски податоци.

Сервисот користи динамички елементи како што се графикони за цените на акциите, табели со сигнали за тргување и листи со анализирани вести. Податоците се преземаат преку API повици од останатите сервиси во системот, како Data Analysis Service и Data Storage Service, и се прикажуваат во реално време.

Корисниците можат да изберат компании преку интерактивно мени, кое автоматски ги ажурира сите визуелизации според избраниот код.

Овој сервис игра клучна улога во системот, бидејќи овозможува крајните корисници лесно да ги интерпретираат аналитичките резултати и да донесат информирани одлуки врз основа на обработените финансиски податоци.

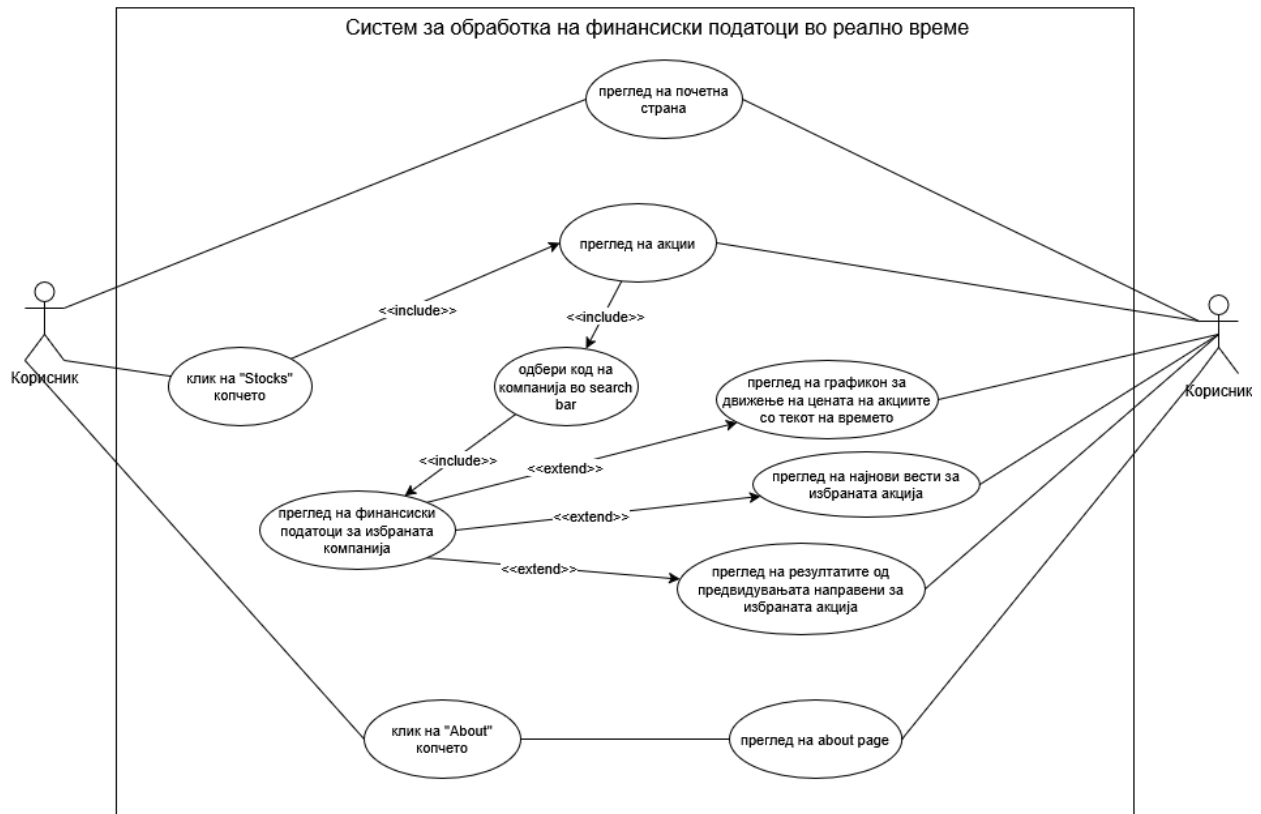
Функциски барања

1. Системот ќе овозможи интеграција со различни надворешни извори преку API за автоматско прибирање релевантни финансиски податоци.
2. Системот ќе овозможи автоматско преземање на финансиски податоци од различни извори.
3. Системот ќе овозможи редовно ажурирање на финансиските податоци.
4. Системот ќе овозможи правилно форматирање на прибраните податоци пред нивна обработка.
5. Системот ќе обезбеди сигурно складирање на обработените податоци во бази на податоци, овозможувајќи лесен пристап до историски и тековни информации.
6. Системот ќе овозможи обработка и анализа на податоците во реално време.
7. Системот ќе користи модели од машинско учење за предвидување трендови и генерирање препораки за тргување.
8. Системот ќе комбинира историски финансиски податоци со нови информации за да овозможи длабинска анализа и предвидувања.
9. Системот ќе обезбеди интуитивен интерфејс кој ќе овозможи лесна навигација, визуелизација и интеракција со аналитичките податоци.

Нефункционални барања

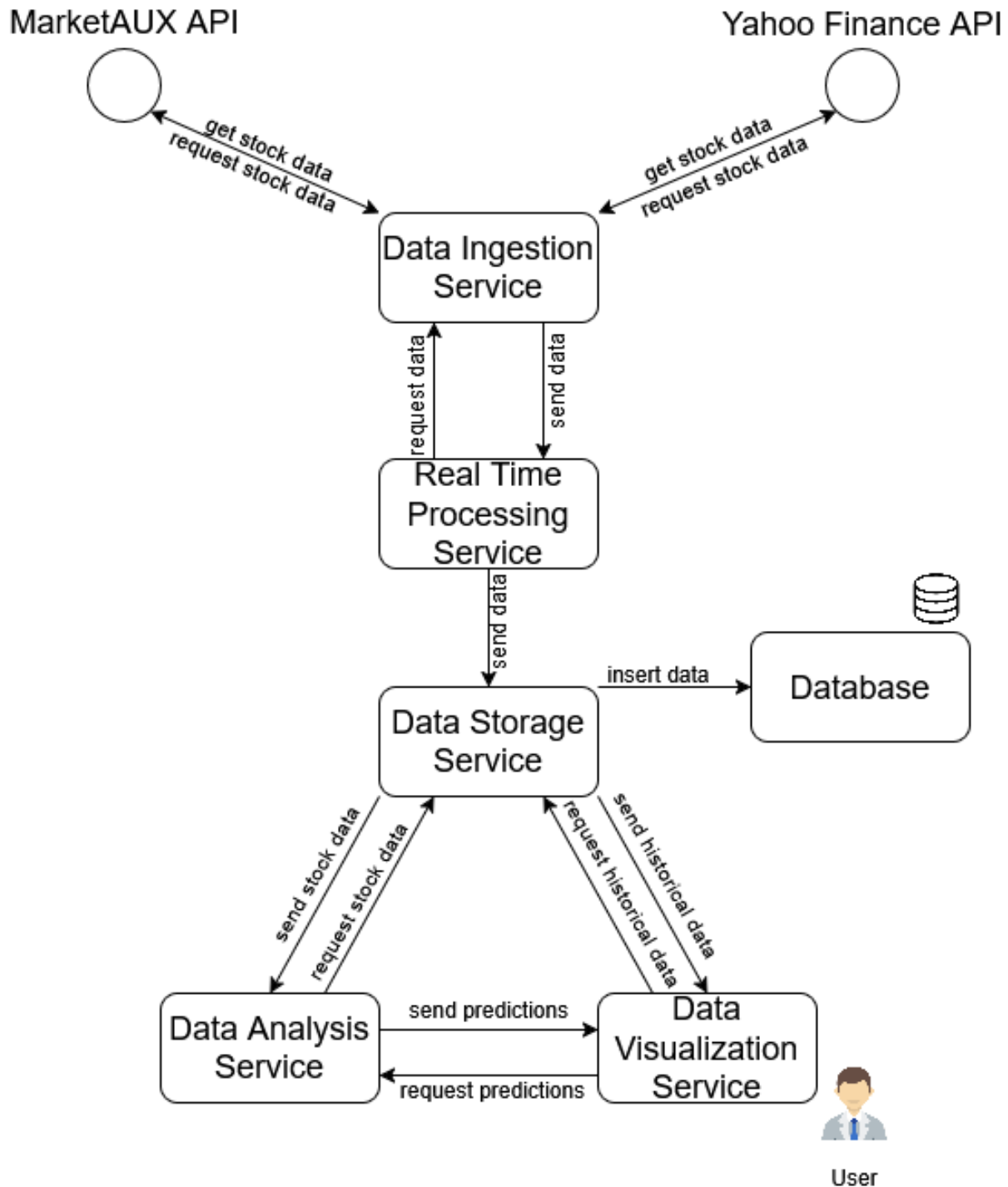
1. Системот ќе биде оптимизиран за максимална брзина и ефикасност.
2. Системот ќе биде способен да обработува големи колочини на податоци.
3. Системот ќе биде модуларен и лесен за одржување.
4. Системот ќе биде достапен 99% од времето.
5. Системот ќе биде достапен на различни уреди и оперативни системи.
6. Системот ќе овозможи лесен пристап до податоците за анализа.

Use Case

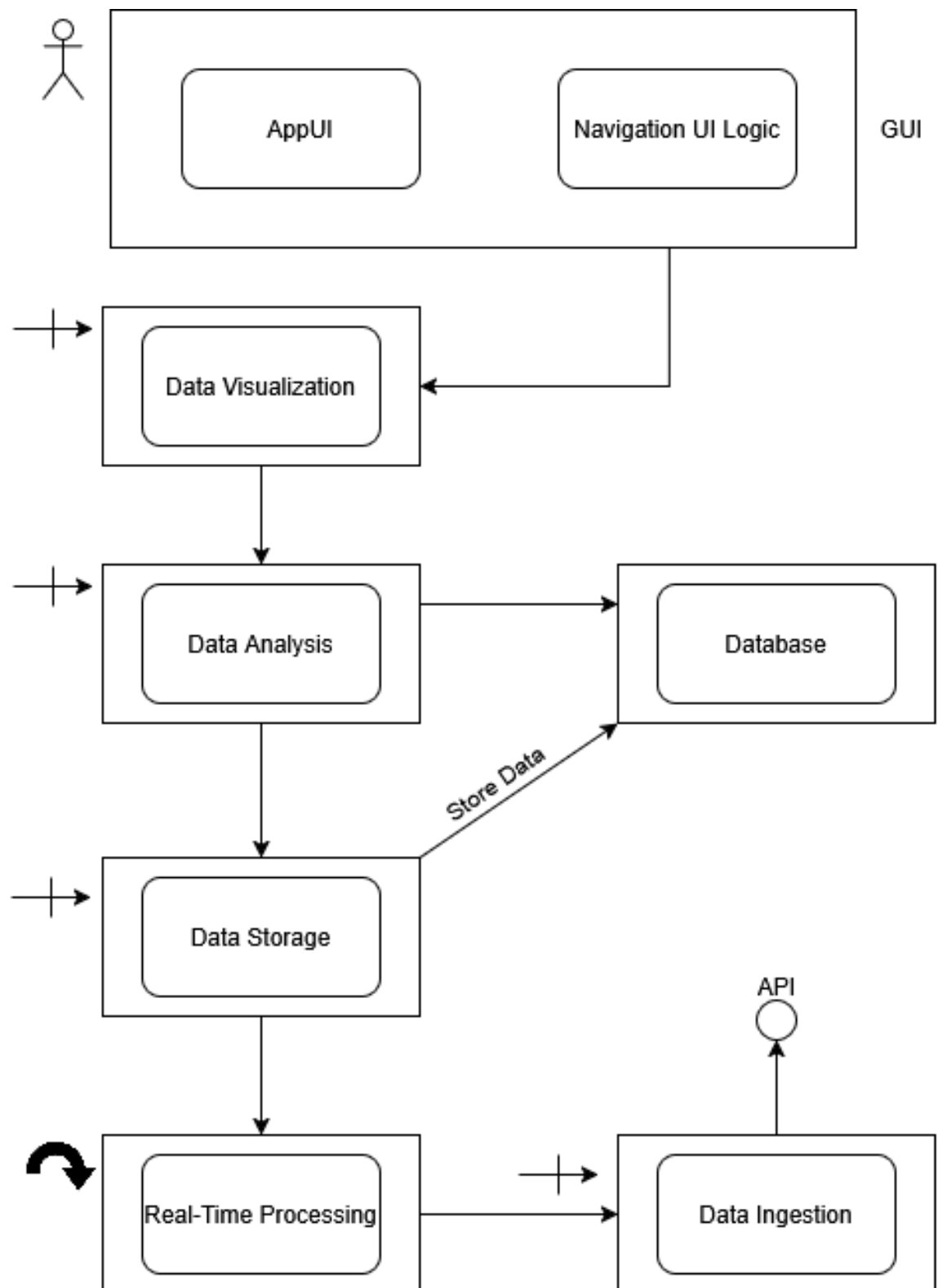


Архитектура на апликацијата

1. Концептуална архитектура



2. Извршна архитектура



3. Имплементациска архитектура

