





MAX K

 github.com/data-silence

 linkedin.com/in/data-silence

 data-silence

 enjoy@data-silence.xyz

Profile

- A financier with 20 years of experience in banks and investment companies who changed careers three years ago
- Made MVP ready ML and NLP projects that is being maintained and developed, that you can use right now
- Two years of ML developing new tools for media and their readers to change and improve the way media work
- Helping educational and media projects
- Looking for a job: living in Turkey, ready for relocation, remote work or startups

Education

Higher School Of Economics, CS faculty

Apr 2021 – Aug 2022

Diploma (excellent): Developing a News Aggregator Using Natural Language Processing Methods

Moscow

Technical Skills

Machine Learning: Scikit-learn, Pytorch, Pandas, Numpy

Data Engineering: PostgreSQL, Docker, FastApi, Pydantic, Streamlit

Data Analytics: BI systems (Superset, Redash, Power BI), A/B testing, Product Metrics

Programming: Python, SQL, Git, Linux Shell

Projects

Research and presentation of bank customers' behavior [git, streamlit web app] :

December 2023

The project is designed to demonstrate some skills in data-science, engineering and analysis: conducting Exploratory Data Analysis, building predictive ML models, formalizing the results of the research in the form of a web service using FastApi, Postgresql and Streamlit.

Python, PostgreSQL, FastApi, Pydantic, SQLalchemy, Streamlit, Loguru, Scikit-learn, Pandas

antiSMI-Project [git, Aggregator bot, Timemachine Bot] :

May 2022 - Present

The Project at the intersection of ML and journalism, which allows using ML to analyze changes in the news flow in real time. Everything is containerized, running remotely on servers. Consists of 4 large independent parts:

- 1. Bot** | *Python, Requests, BS4, PostgreSQL, Docker, SQLalchemy, Pydantic, Loguru, mBart, ruT5, Fasttext, Aiogram, Scikit-learn*
 - * it's a telegram interface to personal and smart news aggregator of russian-language media
 - * uses ML to solve the problem of classification, news clustering, neutral headline and summary generation
 - * removes news garbage (ads, disclaimers, etc.)
 - * combines similar news into stories: minimizes viewing of duplicate news stories
- 2. Timemachine** | *the same stack*
 - * It's a telegram bot to get picture of any day in the past 25 years in Russia
 - * realized news search by arbitrary query using cosine similarity and vector database
- 3. Collector** | *Python, Requests, BS4, PostgreSQL, Docker, SQLalchemy, Pydantic, Loguru, mBart, ruT5, Fasttext*
 - * it's designed to permanently collect fresh news from 50 various sources, process and store them for further use within the Project by other's parts: Bot and Observer
 - * implemented OOP practices to create and storage news data
 - * collected data is validated using Pydantic
 - * trained machine learning models are used to categorize the news and create its summary and title
- 4. Observer** | *+ Apache SuperSet, Natasha, NLP experiments*
 - * is a news visualization and exploration system using Collector databases, SQL and Apache SuperSet
 - * helps to research social trends, create analytical dashboards and build NLP models

Activities

Training

Spring 2022 – Present

You can find most of my studies activities, diplomas and certificates for last three years here

Chocolate Kaggle Competition

May 2023

First place among 55 participants

Experience (IT related):

Sep 2001 – Jan 2022

- * Startup experience (a bank branch was opened)
- * Analysis of big and complex data (risk manager and credit officer of enterprises)
- * Process and people management skills (project manager, deputy director of the bank's branches, member of board of directors)
- * Understanding of budgeting processes, knowledge of business needs