MAHSA ABEDI

LinkedIn: Mahsa Abedi LinkedIn Projects Portfolio: Mahsa Abedi GitHub Personal Website: Mahsa Abedi Website

PROFESSIONAL SUMMARY

Results-driven Data Scientist with a master's degree in AI and 5+ years of experience leveraging AI-based solutions to drive strategic decision-making and fuel organizational growth. Proficient in Python programming and well-versed in various ML and Statistical algorithms, including supervised and unsupervised learning, deep learning, ensemble methods and NLP. Working knowledge of AWS services and Microsoft Certified in developing data pipelines and ML/AI models in Azure cloud. Solid experience in analytical skills including modeling, interpreting, and visualizing large datasets using Tableau, Power BI, and Python libraries.

TECHNICAL SKILLS

- Expertise in Python programming and hands-on experience of various ML and Statistical algorithms, such as supervised learning, unsupervised learning, deep learning using various ML frameworks such as TensorFlow, Keras, Scikit-learn, Pandas, NumPy.
- Expertise in Machine learning techniques including regression analysis, neural network, timeseries analysis, decision tree, and ensemble methods like random forest, XGBoost ,and AdaBoost.
- Microsoft Certified in developing data pipelines and ML/AI models in Azure cloud environment.
- Familiar with utilizing LLM libraries, including LlamaIndex and LangChain to extract valuable insights from text data, optimize machine learning models, and drive data-informed decisions
- Skilled in developing complex SQL queries using Spark-SQL and PySpark in Databricks to extract data from large and complex datasets in snowflake.
- Skilled in advanced SQL with the ability to aggregate and manipulate data through complex queries, views, and stored procedures.
- Expertise in designing advanced dashboards and reports for streaming and historical data using Tableau, Power BI, and Python visualization libraries like Matplotlib, Seaborn, and Plotly.
- Working knowledge of services in AWS like Lambda, S3, EC2, SageMaker, DynamoDB, Redshift, and Kinesis.
- Expertise in advanced statistical methods including Hypothesis Testing, A/B Testing, PCA and Timeseries analysis.

WORK EXPERIENCE

Data Analyst, Analytics and Insights, AI Products

Apr 2022 – July 2023

The Globe and Mail — Toronto

- Developed Spark applications using Spark–SQL and Pyspark in Databricks for data extraction, transformation, and aggregation from snowflake to build an ML Archive model based on the article ages.
- Implemented a sentiment analysis using BERT and word cloud model for news headlines to analyze the prevalence of paywalls and advertisements revenue across topics.
- Built statistical model to forecast optimal stop rate metrics based on revenue and conversion balance, therefore developed strategies to retain registered users on the website.
- Collaborated in building a Conversion prediction model aim to forecast which users are most likely to subscribe or convert from free to paid accounts.
- Produced predictive and perspective A/B testing revenue reports and key performance metrics like confidence level, p-value to analyze revenue opportunities and process gaps using SQL, Tableau Flows, built kite ETL.
- Performed Time-to-Conversion model and Survival analysis like regression to measure the funnel which takes for a user to convert from a free user to a paying subscriber.
- Developed complex SQL queries using spark SQL to analyze large onboarding clients' datasets, resulting in actionable insights and improved business outcomes.
- Conducted data profiling and ad hoc analysis using SQL, Python, and Tableau to identify data trends and patterns of a likely subscriber users.
- Developed automated python and SQL based quality models to ensure the integrity of onboarding and existing clients' data.
- Provided actionable ad hoc analytics insights by visualizing key insights and trends, enabling business stakeholders and management to make data-driven decisions.

Application Data Scientist
Jun 2021 – Apr 2022

MUUTAA
Montreal

- Developed an ARIMA (AutoRegressive Integrated Moving Average) or SARIMA (Seasonal Autoregressive Integrated Moving-Average) to forecast stockout for inventory management of a healthcare company.
- Built an advanced LSTM model for safety stock to forecast demand and inform purchasing decisions.
- Constructed Predictive Modeling Techniques such as Random Forest and Gradient Boosted Model inventory levels and purchasing trends.
- Developed NLP sentiment analysis models using BERT to extract insights from individual clinicians' comments to inform supply chain decisions and align them with clinical practices.
- Built complex SQL stored procedures using PostgreSQL to calculate different statistical algorithms, such as calculating the inventories, stockout, demands, etc.
- Constructed Prophet in time series forecasting models to predict future demand, enabling efficient inventory planning considering multiple seasonality and can capture effects of holidays and special events.
- Conducted data analysis of large data stored in AWS S3 and extracted insights using Athena, resulting in identifying data requirements and improved decision-making.
- Implemented ETL data pipelines to enhance business processes and automate data extraction from various sources into the PostgreSQL database.
- Conducted exploratory data analysis and statistical methodologies such as PCA to identify patterns and trends in data behavior, leading to new features and improvements.

Artificial intelligence Researcher
Aug 2020 – Feb 2023

School of Electrical and Computer Engineering | R&D group
Tehran

- Implemented state of art Noise Reduction in signals using Fully Convolutional Autoencoders and Recurrent LSTM
- Built CNNs to automatically learn hierarchical features from ECG signals and perform heart disease diagnosis classification and arrhythmia classification tasks.
- Conducted analysis on assessing patient disease and applied patient segmentation with clustering algorithms such as K-Means and hierarchical clustering.
- Developed ensemble methods, specifically XGBoost and Adaboost, to enhance accuracy and robustness in ECG classification, resulting in improved diagnostic performance for healthcare applications.
- Implemented Machine Learning algorithms for dynamic branch prediction, resulting in a significant doubling of accuracy in software networks.
- Designed, built, and deployed a set of python modeling APIs for patients' analytics, which integrating AI systems with existing application for various patient disease prediction.
- Prototyped and iterated potential algorithms and pipelines, leveraging historical data to enhance model accuracy and optimize data-driven decision-making processes.

Python Developer
Oct 2018- Apr 2021

Telecom Inc.
Isfahan

- Developed, and deployed a web-based project to provide internet clients with login/registration functionality, account management features, and the ability to modify internet services.
- Created and implemented a REST API for mobile users, facilitating seamless access to retrieve information on available internet plans.
- Contributed to the design and implementation of the database, including database schema development and data migration processes.
- Developed user-friendly and interactive web pages as the frontend component of the web application, utilizing web technologies such as HTML, JavaScript, React, and CSS.

CERTIFICATIONS

- Microsoft Certified: Azure AI Fundamentals 2023
- Career Essentials in Generative AI by Microsoft and LinkedIn 2023
- Microsoft Learn AI Skills Challenge 2023

EDUCATION

M.sc of Computer Science, University of Tehran, Tehran
Sep 2019 – Mar 2022

Teacher Assistant: Tutored undergrad and graduate students in Neural Network, Machine Learning, Statistics course.