Ranjan Ravi

ranjan7796@gmail.com | LinkedIn | Portfolio | GitHub | (813) 729-1906 | Tampa, Florida

PROFESSIONAL SUMMARY

Experienced Data Scientist with a demonstrated history of working in the information technology and services industry. Skilled in Python, SQL, R, Machine Learning, Natural Language Processing, and Data Analysis.

EDUCATION

Master's in Business Analytics and Information Systems, University of South Florida, Tampa, Florida   Expected May 2023

Bachelor's in Computer Science and Engineering, Madras Institute of Technology, Anna University, India   April 2018

WORK EXPERIENCE

Data Science Intern, Bayer, New Jersey   June 2022 – Dec 2022

- Worked on a semantic search tool where users can upload and retrieve information from their corpus.
- Implemented a proprietary scoring algorithm for ranking search results taking into account cosine similarity scores, fuzzy-matching distances, and other metrics.
- Helped in deploying the machine-learning models (T5-based, BERT based) on Amazon Sagemaker.
- Written unit and integration tests for the API using pytest.

Research Assistant and Tableau Mentor, University of South Florida, Tampa, Florida   Nov 2021 – June 2022

- Involved in ongoing research of parsing student resumes using Named Entity Recognition.
- Built our own custom NER model to extract domain-specific entities such as University, Company, Skills, and Years of experience.
- Developed a spacy annotator to generate training data with the spacy annotated data format.
- Worked as a Teaching Assistant for Undergraduate students to help them learn the importance of data-driven decisions and solutions using Tableau.

Data Scientist, TRIMBLE Inc, Chennai, India   Nov 2020 – Aug 2021

- Involved in Data Collection, Data Cleansing, and loading the project details into Snowflake tables using Python.
- Experimented with different time series forecasting algorithms in predicting the future revenues of the Civil Construction division of Trimble and built an efficient model (Facebook Prophet) with 83% accuracy.
- Built an ML pipeline by deploying the model in AWS Sagemaker and wrote its inference function in AWS Lambda.
- Developed dashboards with monthly and quarterly revenue predictions using Tableau.

Associate Software Engineer, BNP PARIBAS India Solutions Private Limited, Chennai, India   Jun 2018 – Nov 2020

- Involved in an Agile software design and development of Reporting platform for generating tax reports for clients residing in more than 100 countries.
- Written PL/SQL procedures and SQL scripts for implementing the tax calculation based on the Double Tax Treaty as per the market requirements.

TECHNICAL SKILLS

Programming Lang. & Frameworks:	Python, R, Flask, FastAPI, PL/SQL, Java, Dash, HTML & CSS
Databases:	Oracle, PostgreSQL, MS SQL, SQLite
Machine Learning Models:	NLP models, SVM, KNN, Ensembles, Neural Networks (CNN, RNN)
Cloud Technologies:	Amazon web services - EC2, S3, Sagemaker, Lambda, Snowflake, Azure ML Studio
Python Packages:	TensorFlow, Scikit-Learn, NLTK, SpaCy, Pandas, Numpy, Beautiful Soup, OpenCV
Reporting Tools & packages:	Tableau, Matplotlib, Seaborn, Plotly, Jasper reports, PowerBI
Time Series Forecasting algorithms:	Facebook Prophet, Neural Prophet, ARIMA, Exponential Smoothing, LSTM

PROJECTS

- **NLP Albumentation:** Experimented with different data augmentation techniques to text data. Examined whether the NLP albumentation helps improve the classification model's performance. [GitHub](#)
- **Text Analytics:** Analyzed the inaugural speeches of all the US presidents using NLP techniques. Performed Topic modeling in finding the common theme of each of the president's speeches. [GitHub](#)
- **Covid-19 Tracker**: Built a website with an interactive chart to display the trend of confirmed, recovered, and deceased coronavirus cases on a per-day basis using the Python Dash framework. | [URL](#) | [GitHub](#)
- **Sentiment Analyzer:** A big data project to capture the sentiment of E-Commerce user reviews stored in a Hadoop File System using Apache PySpark and MLlib. [GitHub](#)
- **GitHub Analytics:** Built a Classifier to identify potential candidates based on their GitHub profiles in order to minimize the time and effort the recruiters put in during the recruitment process. | [GitHub](#)
- **Linear Regression Model and Web Scrapping using Real Estate Data:** Designed and built a web scraper program to scrape house data from a real estate website and built a simple Linear Regression model | [GitHub](#)
- **Personal Portfolio:** Developed a compatible personal portfolio using HTML5/CSS3 and Bootstrap, and deployed it in the Heroku Cloud Platform. | [URL](#) | [GitHub](#)