Rajdeep Biswas

Data Scientist / Machine Learning Engineer

Address Malbazar, India, 735221
Phone +91 9874 175 879
E-mail r4jdeepbiswas@gmail.com
Website https://rajdeep-biswas.github.io
GitHub https://github.com/rajdeep-biswas
LinkedIn https://linkedin.com/in/rajdeep-biswas

With a year of demonstrated ability to deliver reliable ML / DL applications working in R&D projects of a major product based company, and a background of high-quality independent research at college and company levels surrounding the areas of Time Series Analysis and Natural Language Processing, I wish to continue to push all boundaries in these areas and come up with innovative solutions via rigorous research and application.

Skills

Languages / Libraries: Python (Numpy, Pandas, Sklearn, Matplotlib), Tensorflow (Keras), PyTorch, MATLAB / GNU Octave, Flask, C / Java.

Areas: Sequence2Sequence Natural Language Processing Models, Transformers, Unsupervised / Semi-Supervised Anomaly Detection, LSTM Autoencoders, Deep Learning, Machine Learning, Data Engineering, NLP Multilabel Classification, RESTful Microservices, MVC Architecture, Backend Engineering & Model Deployment.

Work History

August, 2020 - Present

Engineer

SAP Labs, Bangalore

Team: **Innovation Center Network** (Oct 2021 - Present)
• Presently at the feasibility check stage of an **NL2SQL** product. Assessing various pretrained models like GPT-J, T5 + WikiSQL, SQLNet, etc. on a limited training dataset.

Team: **API Management** (June - Dec 2021)
• Developed an **Anomaly Detection** service for API traffic across various Tenants. Wrote a Deep Learning model that uses an **LSTM Autoencoder** architecture. This used Tensorflow (Keras). Attained ~90% accuracy over 1 year of training data.
• Analyzed decomposition based libraries like Facebook-Prophet and Microsoft's SR-CNN. Containerized all three models via Docker carrying RESTful Flask applications, for an ensemble consensus-based outlier detection service.

Hackathon: **InnVent** (Sept - Oct 2021)
• Wrote a **Custom Transformer Sequence-to-Sequence** architecture for an **NL to Cloud CLI Commands Assistant** bot. Used PyTorch. Attained ~80% accuracy, on a dataset that was hand-curated majorly by myself from web resources..

Hackathon: **Innovision** (May - June 2021)
• Created an NLP **Multilabel Classification** pipeline which used a **Bidirectional LSTM** architecture using Tensorflow (Keras) in order to identify the type of service required based on a given sentence and predict a class.

Team: **Spend Visibility, Ariba** (Sept 2020 - Apr 2021)
• Implemented a feature enhancement that significantly minimized pipeline freezes, from **90%** of the time to **< 10%** of the time, which were caused by dataload limitations involving sizes of 1m+ of records. (SpringBoot, Python)
• Took over various product bugfixes and delivered them in record time. (SpringBoot, Angular, Python)
• Tested various Time Series forecasting methods that SAP HANA Db PAL (Predictive Analysis Library) had to offer, wrote a custom Python pipeline that analyzes about eight of said algorithms and asses the performance of each across the entire dataset containing various customers, commodities and their expenses.
Post analysis of best algorithms, implemented an expense forecast Proof of Concept with custom stored procedures written on HANA invoked via a SpringBoot backend.

June - August, 2019

Intern - Conversational AI

SAP Labs, Bangalore

• For **CoPilot,** assessed the performance of various ML clustering and classification algorithms on matching intents and contributed to its NLP pipeline.
• Took up various bugfixes on the frontend of the product, based on ReactJS. Responded to and fixed multiple internal and customer incidents and backlogs.
• Contributed in migration of a few modules of the product's backend from Ruby to Java SpringBoot, a microservices framework.

Publications

• COVID-19 Prediction Effectiveness Time Series and Lockdown. **International Conference on Computational Intelligence in Data Mining (ICCIDM) 2021**. (Submission concluded. Presentation incoming in Dec, 2021).
• Predictive Analysis of the Recovery Rate From Coronavirus. **International Conference on Cyber Intelligence and Information Retrieval (CIIR), 2020**. (Done in collaboration with Master's seniors during final year of college).

Education

2017 - 2020

Bachelor of Computer Application

Institute of Engineering & Management. 8.89 DGPA.