Anand Krishnamoorthy

Krishnamoorthy.a@northeastern.edu | 857-800-3336
https://www.linkedin.com/in/anand-krish | AnandKrishnamoorthy1 (github.com)
Portfolio website: https://anandkrishnamoorthy1.github.io/

An enthusiastic Data Scientist with a passion to keep improving (and learning) and an ambition to develop applications that will revolutionize and create an impact.

EDUCATION

Northeastern University, Khoury College of Computer Science, Boston, MA — Sept 2019 – Dec 2021
- Master of Science in Data Science, GPA: 3.7

Amrita School of Engineering, Kollam, India — Aug 2011 - June 2015
- Bachelor of Science in Mechanical Engineering

TECHNICAL KNOWLEDGE

Tools: Python, Alteryx, PyTorch, Keras, SQL, Informatica, Tableau and Excel
Databases: MySQL, Oracle 11g (SQL, PL/SQL)
Technical skills: Data preparation, project brainstorming, Regression, Classification, Clustering, Genetic Algorithm, NLP, Text Similarity, Text Summarization, Named Entity Recognition, SQL, and ETL

YES BANK GLOBAL DATATHON : Data Science Competition Winner

Ticket Automator Solution — Oct 2018 - Jan 2019
- Contested and **won** the Yes Bank global datathon (**6000+ participants**) for developing the Ticket Automator Solution (TAS), which automates customer logs, prioritizes the logs, and highlights the key phrases.
- Received **citations** of the Ticket Automator Solution (TAS) in technical blogs. Solution classifies customer logs-using classification techniques, highlights key phrases-using RAKE, and uses sentiment of the text to assign priority.

PROFESSIONAL EXPERIENCE

Audax Private Equity, Boston, MA
Data Science Intern — Jan 2021 – Jun 2021
- Worked on RFM analysis along with clustering and data analysis to understand company's **market share** and **customer segmentation** for an e-commerce company.
- Initiated the company's first data science project to predict the next equipment repair for a portfolio company. Developed a tree-based **regression** model with confidence intervals (upper and lower bounds) for the prediction. The model predicts the next equipment repair (+/- 30 days) with **75% accuracy**.

Tata Consultancy Services, Chennai, India
Systems Engineer/Data Scientist — Mar 2016 – Jul 2019
- Built a software IDEA, which reads unstructured data (like PDFs) and **extracts information**. IDEA classifies text, extracts tables, and summarizes text. IDEA also uses spacy to parse Named entities (NER) and their dependencies.
- Successfully developed an **NLP** solution that classifies and suggests resolution for service tickets, which can drastically reduce manual effort **by 30%.** Doc2vec model and random forest model were utilized as they performed better on classification accuracy.
- Built **ETL** pipelines and helped build Smart Mapper, an **SQL** solution, which maps source attributes to target attributes in a data model. Reduces the manual effort of Data Modelers by 15%.
- Developed a **Log Analytics** POC using python, HDFS, and HIVE which can extract job execution and failure statistics of weekly and monthly jobs.

ACADEMIC PROJECTS

CommonLit-Readability-Prize (CLRP)
Kaggle Competition — Jul 2021 – Aug 2021
- Fine-tuned **BERT and RobERTa** (SOTA language models) with the CLRP text and ensembled the models to accurately predict the reading difficulty of the given text to achieve an RMSE score of 0.461. In comparison, the winning submission had an RMSE score of 0.446.

Which celebrity do you resemble?
Harvard University, Cambridge, MA — Jun 2020 – Aug 2020
- Developed a Deep learning model with a CNN architecture (VGGFace) by employing data augmentation, transfer learning and Face detection techniques to recognize celebrity faces with an accuracy of 86% accuracy.
- Resulted in **5-8% higher accuracy** compared to several other implementations found online.