ANAND KRISHNAMOORTHY

46 Westland Avenue, Boston, MA 02115 |Krishnamoorthy.a@northeastern.edu | 857-800-3336
LinkedIn: https://www.linkedin.com/in/anand-krishnamoorthy-183213a5
Portfolio website: https://anandkrishnamoorthy1.github.io/
Available: Jan 2022

EDUCATION

Northeastern University, Khoury College of Computer Science, Boston, MA — Sept 2019 – present
Candidate for a Master of Science in Data Science, GPA: 3.6 — **Expected Graduation:** Dec 2021

- Courses: Linear algebra and probability, Algorithms, Natural Language Processing, Supervised Machine Learning, Introduction to Deep Learning, Unsupervised Machine Learning, Reinforcement Learning

Amrita School of Engineering, Kollam, India — Aug 2011 - June 2015

- Bachelor of Science in Engineering
- Relevant Coursework: Calculus, Statistics, Differential Equations, Operations Research, Computer Programming

TECHNICAL KNOWLEDGE

Tools: Python, Alteryx, PyTorch, Keras, SQL, Informatica, Hive, Hadoop, Tableau and Excel
Databases: MySQL, Oracle 11g (SQL, PL/SQL)
Technical skills: Regression, Classification, Face detection, Clustering, Genetic Algorithm, GANs, Text Similarity, Text Summarization, Named Entity Recognition, Oracle DBs, 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 other techniques to segment customers for an e-commerce company.
- Developed a tree-based regression model with confidence intervals (upper and lower bounds) to predict the next equipment repair for a portfolio company. The project also required brainstorming ideas and communicating the results to stakeholders. The model accurately 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 takes regulatory documents and segregates requirement and informative sentences. IDEA extracts tables using PDF parser and summarizes text using a scoring function (and by genism module) by identifying action verbs, POS tags. 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 efforts and customer care/support expenses. Doc2vec model and random forest model were utilized as they performed better on classification accuracy.
- Contributed to the development of various POCs, ETL pipelines and helped build Smart Mapper, a SQL solution, which maps source attributes to target attributes in a data model. Reduces the manual effort of Data Modelers.

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. 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 33 celebrity faces with an accuracy of 86% accuracy.
- Resulted in 5-8% higher accuracy compared to several other implementations found online.