Hasan Khan

MSCS Graduate @ NYU Courant

hk3550@nyu.edu
+1-804-980-5375
theotherkhan.github.io

Professional Experience

NLP Engineer – Healthcare Risk Advisors
6/2022 – 1/2023
5/2023 – Current

Implementing NLP pipelines utilizing large language models (LLMs) for text classification tasks on medical claims data. Exploring supplementary training techniques for LLMs. Modeling using PySpark and SparkNLP in Databricks.

Data Scientist – LMI, U.S Federal Consulting
8/2018 – 5/2021

Provided advanced analytics and software development services for federal agencies, with a focus on building machine learning capabilities. Highlighted work:

Inventory Management - Built supervised regression models to optimize rare inventory lead time estimation for a U.S Defense Agency. Random forests trained on 10+ years of inventory procurement data resulted in a 37% improvement of lead time predictions for rare procurements over baseline metrics, with an annual expected savings of $45M on inventory holding costs.

Survey Analytics - Designed and developed a survey analytics web application for NASA Human Resources. Utilized BERT language models for sentiment analysis and topic modeling on free response survey comments. Integrated models into Python Dash webapp and visualized sentiment and topic classification outputs. Deployed webapp on Shiny.

Research

Improving Knowledge Distillation
1/2023 – Current

Current research on improving the efficacy of Knowledge Distillation in LLMs by incorporating novel curriculum learning methods, where student models selectively rely on teacher distillation loss depending on datapoint difficulty. Conducted as part of ML and Computational Linguistics coursework.

Neuronal Modeling – Meliza Lab, UVA Psychology Department
5/2016 – 8/2018

Modeled neurons in Python to simulate auditory cognition processes in Zebra Finches. Coded Recursive Bayesian Filters (RBF) for biophysical parameter estimation of High Vocal Center neurons. Classified auditory neurons into subtypes based on ion channel makeup.

Education

New York Universty – M.S Computer Science
9/2021 – 5/2023

GPA: 3.4

University of Virginia – B.A Computer Science, B.A Cognitive Science
8/2014 – 5/2018

GPA: 3.3

Proficiencies

Languages: Python, Java, C++, SQL, HTML, Ada
Libraries: sklearn, keras, PyTorch, PySpark, MLflow, SparkNLP, HuggingFace
Platforms: AWS, Git, Docker, Databricks
AI/ML Coursework: ML, NLU, Computational Linguistics, ML for Healthcare
Clearances: US Federal - Secret