NAFIS ABRAR

https://nafisabrar.com
GitHub https://github.com/nafabrar
+1 778-929-4772
nafis.abrar@alumni.ubc.ca

WORK

Jan 2020-Present
Django,HTML/CSS/ js, Pytorch/Fast.ai/ Tensorflow

ML RESEARCH ASSISTANT
Vancouver

at Laboratory for Computational Intelligence (LCI) (UBC CS)

- Currently working on a research paper on automatic essay grading using transformer language models such as BERT,ROBERTA,Electra.
- Implemented NLP pipeline for grading ASAP essay dataset which resulted in a 9% increase from current state of art Transformer model.
- Developed automatic tracking features using javascript and django to collect data on Mechanical TA peer grading software.

Jan 2020-Present
Node.js, Firebase, React, HTML/ CSS,Python ML libraries

CO-FOUNDER/ SOFTWARE ENGINEER
Vancouver

at Edutechs.org

- Co-founded edutech - the first online multi-video classroom software in Bangladesh - which is being used by tutors.
- Implemented image to text functionality. Currently working on language models for automatic grading and plagiarism detection.

Jan 2019 - Aug
Scikit-learn, XGBoost,SQL, PySpark, Docker, MS Azure

DATA SCIENTIST
Toronto

at ScotiaBank(Artificial Intelligence and Machine Learning Team)

- Worked in an Agile team for Scotiabank's global fraud detection AI software. Applied ensemble methods such as XGBoost for prediction and inference.
- Performed sensitivity analysis and hyperparameter tuning to improve model performance by 4%.
- Designed and code reviewed a NLP model which accurately monitored customer consent from text transcriptions. This model improved accuracy by 43% from the previous version.

Jan 2018 - Aug
Django,Pandas,Scikit-learn,PostgreSQL HTML/CSS/JS,MS Azure,Github,Linux

SCIENTIFIC SOFTWARE DEVELOPER
Vancouver

at BC Cancer Research Centre (Sohrab Shah Lab)

- Worked on a machine learning research project for integrating genomic data with imaging data of cancer cells to classify dead/alive cells. The classifier resulted in an prediction accuracy of 84%.
- Performed data analysis and implemented machine learning algorithms for cancer cell clustering problems.
- Implemented, extended and documented python APIs and REST interfaces.

EDUCATION

May 2020

B.Sc. in Computer Science
Vancouver

at University of British Columbia

Focus on Machine Learning
Notable courses: Intelligent Systems/AI, Machine Learning, Advanced Database, Computer Vision, Software Engineering, Statistical Learning.

RELEVANT PROJECTS

June 2018

PIMS BC DATA SCIENCE NLP CAPSTONE PROJECT

Language: Python Frameworks: Pandas, Scikit Learn,TextBlob, spaCy

- Worked in a team of 10 to determine intent and create knowledge base from live chat transcripts. The data set was provided by Comm100 which includes online chat sessions.
- The goal of the project was to cluster/correlate chat sessions and build a knowledge base in an automated way using mathematical models which we successfully achieved.

Publication

- Scalable whole genome sequencing of 40,000 single cells identifies stochastic aneuploidies, genome replication states and clonal repertoires. *
- Clonal Decomposition and DNA Replication States Defined by Scaled Single-Cell Genome Sequencing. **

* https://www.biorxiv.org/content/early/2018/09/13/411058
** https://www.sciencedirect.com/science/article/pii/S0092867419311766

Programming Languages

- Python — Competent (2 years)
  Experience : Scikit-learn, Machine learning algorithms, Django- UBC EOSC website, Rhoads, NLP
- Java — Competent (1 year)
  Experience : Restaurant App, Advanced Calculator, DNS Server, FTP Client, Gym Database using JDBC
- Unix/bash — Competent (2 years)
- SQL — Competent (1 year)
  Experience : GYM Database, UBC EOSC website
- C — Basic (8 months)
  Experience : FTP Server , x86 implementation

Web development and Design

- HTML and CSS - Basic (4 months)
  Projects : UBC Eosc website, Rhoads website
- TypeScript - Basic (3 months)
  Projects : Insight UBC (course project)
- Django - Proficient (8 months)
  Projects : UBC EOSC website, Rhoads website

Machine Learning

- Python - scikit-learn, Pandas, matplotlib
- PySpark
- Hadoop
- TensorFlow*
- Pytorch/Fast.ai
- MXnet*

*Currently learning for project

Dev-ops

- Github
- Docker
- Jira