GAURAV **MUDBHATKAL**
MANAGEMENT ENGINEERING AT THE UNIVERSITY OF WATERLOO

✉ gmudbhat@uwaterloo.ca 🌐 gaurav613.github.io
📞 (+1)5197292861 📍 University of Waterloo
in linkedin.com/in/gmudbhat ⚙ gaurav613

SKILLS

Programming Languages: Python, Java, Scala, HTML/CSS, R, C++, JavaScript
Data/ML: pandas/NumPy, Spark, scikit-learn, Keras/TensorFlow, NLTK/gensim/spaCy, matplotlib/seaborn, Tableau/Excel
Tools and Technologies: Git, Linux, Airflow, Jenkins, Docker, Jupyter/Zeppelin, Anaconda, Maven, Google Cloud, jFrog
Project Management: Scrum, Problem Analysis/Solving, JIRA, Confluence, Notion, Communication/Reaching Out

WORK EXPERIENCE

Data Engineering Intern · KPMG - Lighthouse Exponential   July 2020 to Aug. 2020 · Remote
- Will be working as a Data Engineering Consultant at KPMG from July

Data Scientist, AML Insights · RBC - Royal Bank of Canada   Sept. 2019 to Dec. 2019 · Toronto, Canada
- Built an in-house capability to gain insights from Suspicious Transaction Reports with Natural Language Processing
- Preprocessed and tokenized text using **NLTK**, generated embeddings for ~10,000 documents by applying **Doc2Vec** algorithms
- Designed a neural network-based multi-label text classifier with **Keras** - attaining accuracy of 96% and a precision score of 82%
- Facilitated the ability to search for reports based on document similarity using **gensim**, performed Entity Recognition with **spaCy**
- Analyzed client behavior by implementing clustering techniques such as **KMeans** and **HDBSCAN** on client data
- Automated package deployment to the **jFrog Artifactory** by integrating with **GitHub** and **Jenkins**

Data Analytics Co-op · BBM - BlackBerry Messenger   Jan. 2019 to Apr. 2019 · Mississauga, Canada
- Simplified data processing in an Apache **Spark** environment with **Scala**- transforming raw data into tables written to **BigQuery**
- Optimized the ETL pipeline by scheduling jobs in Apache **Airflow**
- Investigated sudden spikes in data using **Tableau** and Apache **Zeppelin**
- Enhanced CI/CD of the data pipeline by providing access to **Jenkins** scripts via version control
- Configured dependencies in **Maven** to generate complete documentation of the entire codebase

Junior Developer · Perpetuuiti   May 2018 to Aug. 2018 · Pune, India
- Developed a CLI-based Object Detection application using the Deep Neural Networks module in **OpenCV**
- Gained valuable knowledge in Machine Learning concepts through online courses and material provided by the mentor
- Built, trained and tested Convolutional Neural Networks from scratch using **Keras** and **TensorFlow** to detect objects of classes present in the CIFAR-10 dataset, achieving an accuracy of 92.3%

PROJECTS & EXTRACURRICULARS

Entrepreneurship Co-op - Billion Dollar Problems   May 2020 to June 2020
- Conducted problem analysis and market research on storage/analysis of digital evidence data in Law Enforcement Technology
- Selected to compete in the Quantum Valley Institute's Problem Pitch Competition, for a chance to win up to $15,000

WATonomous - Infrastructure Core Member
- Migration of the development environment from VMs to **Docker** containers
- Visualization of ROS messages(Sensor Data) using RViz

Stock Prediction Model
- Predicted stock prices based on historical data using the concept of **LSTM**(Long Short-Term Memory) with **Keras**

Toxic Comment Classifier
- Detected toxicity level of online comments by implementing a Logistic Regression Pipeline in **scikit-learn**

Personality Matching Application
- Created an application using **Tweepy** to display the top 5 similar traits between two people based on their tweets
- Implemented **IBM Watson's** Personality Insights API to find out and compare the traits of both the personalities

EDUCATION

University of Waterloo
BASc Honors Management Engineering 2022

deeplearning.ai (Offered by Coursera)
Tensorflow in Practice Specialization 2020