KOUSHIK K
SOFTWARE & MACHINE LEARNING

CONTACT

- koushik8.kolli@gmail.com
- +918310696989
- Hyderabad, India
- linkedin.com/in/koushik-kolli
- github.com/koushikkolli

PROFILE

Technology Analyst, working in Deep Learning based project at Infosys

EDUCATION

2012-2016

UNIVERSITY VISVESVARAYA COLLEGE OF ENGINEERING, BENGALURU

Bachelor of Engineering, Mechanical

CERTIFICATIONS

- Certified Data Scientist and AI Expert, Edvancer
- DP-100 certification
- Machine Learning, Coursera
- Deep Learning Specialization
- Machine learning specialization
- Google IT automation certificate
- Natural Language Specialization

TECHNICAL SKILLS

- Python, JavaScript, Java, React, NodeJS, Scala
- Machine Learning and Deep Learning
- TensorFlow, Pytorch
- Predictive Modelling, NLP
- Scikit-learn, Matplotlib
- Hadoop, Spark, PySpark
- RDBMS and MongoDB
- Apache Nifi, Apache Kafka

EXPERIENCE

2016 - PRESENT

Technology Analyst | Infosys

Current Project

- Developing application for migrating database from Oracle to Postgres using **Ora2PG and Deep Learning** based approach. Which already helped in migrating of thousands of databases for various clients and generated over **2 million USD of revenue.**
- For converting database objects and queries using Deep Learning, we are using **Transformer model**. The deep learning model has helped in **reducing the manual work by 20%** on top of Ora2PG.

Past Projects

- Developed application for Data Profiling and Cleaning. Used **random forest and DBSCAN** for filling the missing values and finding the outliers respectively.
- Developed an ETL application for migrating the data between different databases which generated **revenue over 5 million USD**. Implemented bulk data loading into MSSQLServer and Oracle using BCP and SQLLoader.

PERSONAL PROJECTS

- Detecting whether the person is wearing the mask or not and distance between the persons using TFOD. **Created my own data of 1500 pictures** of with and without mask in approximately equal proportions. Labeled the data using LabelMe. Used transfer learning from **Faster RCNN model**. I have directly used a ssd lite model to detect the persons as person is one of the class in the pre-trained ssd lite model and calculated the distance between two persons.

- Building a solution to identify spam questions on Quora based on intention as apparent question text and description. **LSTM model** trained from scratch took too long to fit and did not perform well. Training was made further challenging because of large volume of data when I tried to include description of the questions. Incorporating pertained GLOVE embedding [twitter data with dimension 50 as it aligned best with social media language] reduced training time while **improving performance to 89% accuracy.**