Akansh Katyayan

📍 London, UK   ✉ akanshkatyayan27@gmail.com   📞 +44 7438505052   in linkedin.com/in/akanshkatyayan/

💼 PROFESSIONAL EXPERIENCE

Data Scientist
HSBC Bank plc (Infosys Consultant)
07/2020 – 09/2021
London, UK
- Collaborated with clients to best understand the operational and strategic needs in terms of Operational risk data analysis.
- Achieved 28% increment in the Controls effectiveness by integrating NLP services like NER, and Text Similarity identification.
- Built an NLP system and acquired a 20% increment in the correctness of Ops Risk data by classifying and recommending an appropriate Risk Taxonomy.
- REST APIs development and deployment. Management of Unix/Linux environments. CI/CD Pipelines, Git Version controlling.
- Developed data pipelines in Python for ETL operations including for a dashboard representation of business objects in Qlik.

Data Analyst
HSBC Bank plc (Infosys Consultant)
01/2019 – 06/2020
London, UK
- Information Retrieval from multiple sources with archival process.
- Automated deployment of Fastmaps (OOTB utility in OpenPages) which saved the weekly manual effort of multiple hours. Implemented using Python ETL job, web-scraping, and SSH configurations.
- Slashed manual effort by 80% by automating the process of extracting data through Cognos Analytics, transforming, and validating the data using Python scripts.
- Managed offshore team and conducted Scrums to track sprint progress in an Agile environment. Actively involved in Sprint planning, grooming, retrospective meetings, branching strategies and release management for the project

Associate Consultant
Infosys Ltd. | Client: HSBC Bank
09/2016 – 12/2018
Pune, India
- Handled data migrations, and reference data for ORM. Generated and handled dashboards and dimensional reporting in Cognos Analytics to provide insights into the operational risk data.
- Developed TFUI, Triggers, and Helpers. Involved in handling configurations in IBM OpenPages 8.1.
- Closely worked with the business stakeholders to gather the change requests, and transform them into technical requirements in JIRA.

Programmer Analyst
Cognizant | Client: LPL Financials
05/2014 – 08/2016
Pune, India
- Design and development in cutting-edge ECM technology Documentum for documents metadata.
- Experience in configuration, up-gradation, and administration of Documentum Suite of Products.
- Developed web-based applications using Documentum Foundation Classes (DFC) 6.x/7.x and Java/J2EE, REST APIs implementation. Unit testing, & deployment of the Code.

🎓 EDUCATION

Master of Science - Big Data Science
Queen Mary University of London (Grade: Distinction)
09/2021 – present
London, UK
Coursework: Machine Learning, Natural Language Processing, Data Mining, Applied Statistics, Deep Learning and Neural Networks, Big Data Processing, Cloud Computing, Neural Networks and NLP.

Bachelor of Technology - Information Technology
Rajasthan Technical University (Grade: First Class)
09/2009 – 06/2013
Ajmer, India

📁 PROJECTS

Ethereum Analysis ↗
Analysis of Etherum contracts, transactions, gas, scams, wash trades using Big Data Hadoop MapReduce & Spark
- Evaluated the top 10 smart contracts and miners between 2015-2019 with average transaction values.
- Identified the most lucrative scams, forks wash trades and correlation between gas price change over time for each transaction.

Audio classification (Songs) with MLEnd Hums and Whistle dataset ⧉

Samples: 12000+ Hums and Whistle recordings.

- Machine learning pipelines for the classification of songs based on Hums and Whistle dataset.
- Audio signals were decomposed to extract various features (MFCC, RMS, ZCR,), utilized by multiple machine learning classification models like SVM, kNN, and Neural networks (CNN, MLP), Deep Learning models, Hyperparameters optimization

Modalysis ⧉

Text Summarisation - BERT Summarizer

- Cloud-based Flask application with REST APIs using state-of-the-art Machine Learning model to perform text summarization with a user management system. NLP Model: BERT Summarizer. Deployed on GCP (Google App Engine) and Heroku

CRF Sequence Tagging for Movies Dataset and Vector Space Semantics for Similarity

- Optimization of the performance of CRF tagger for movie trivia questions and answers data using BIO tags, POS tagging, NER, feature extraction and hyperparameters tuning.
- Cosine-Similarity based classification of characters using Vectors semantics. Applying multiple pre- processing techniques, KBest feature selection, Dialog context, N-grams, DictVectorizer, TF-IDF.

Deep Learning projects

- LSTM for Text Classification: RNN, using embeddings layers/pre-trained GloVe Embeddings.
- CNN/BiLSTM performance evaluation for Text Classification and Aspect based Sentiment Analysis.
- NMT with Attention for Machine translation
- Aspect Based Sentiment Analysis using Transfer learning – BERT, Attention.
- MNIST Fashion Dataset classification using MLP, CNN, MXNET ⧉
- Encoder-Decoder Dialogue System. ⧉

🧠 SKILLS

Programming Languages

Python, R, Java, C, C++, Javascript, HTML/CSS

Python/ML Libraries

Pandas, Numpy, Seaborn, Matplotlib, PyTorch, Tensorflow, Keras, Scikit-Learn, XGBoost, OpenCV, PySpark, Pipelines, Beautiful Soup, Requests, PyTest, mrjob, Neural Networks, Time Series Forecasting, NLP models - RNN, LSTM, BioBERT, BERT, CNN, SpaCy, SciSpaCy, NLTK, TabPy, plotly

Cloud Computing

AWS (EC2, S3, RDS), GCP (Compute Engine, Storage, Kubernetes Engine, App Engine)

Database

Oracle DB, SQL Server, MySQL, Amazon Dynamo DB, MongoDB

Big Data

Hadoop/HDFS, MapReduce, Spark, Hive, Mahout, Sqoop, Flume

Tools/IDEs

Tableau, Google Colab, Pycharm, Spyder, RStudio, IBM OpenPages, Cognos, VS Code, Eclipse, JIRA, Confluence, Postman, Diagflow, Putty, Atom

🎖 ADDITIONAL COURSES

Python and Statistics for Financial Analysis

Hong Kong University of Science and Technology

Machine Learning

Stanford University

Google Analytics

Google

🏅 EXTRACURRICULAR ACTIVITIES AND ACHIEVEMENTS

Awards and Recognitions

- *Financial Service (FS) Star* award by Infosys for getting the top rating for 4 consecutive years. **Infosys, 2021**
- *Multiple 'Insta Award'* for DZ metric result – Double the release and Zero incidents/defects. **Infosys 2018-2020**
- 'Instantly Inspired' award from the client (LPL Financials). **Cognizant, 2016**

Positions

- **Sports Officer:** Elected for representing the interests of sports clubs and assisting members of Union staff.
- **Course Representative** in Bachelors – liaison between faculty and students of the IT branch.