Akansh Katyayan

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

💼 PROFESSIONAL EXPERIENCE

Data Scientist
HSBC (Infosys Consultant)
07/2020 – 09/2021
London, UK

- Collaborated with multiple stakeholders to best understand the operational and strategic needs in terms of Operational risk data analysis, NFR Process led data insights.
- Achieved 28% increment in the Controls effectiveness by building Controls recommendation engine for Risks using NLP Text Similarity and search (BERT Embeddings, Milvus vector database). Similar model was utilised to recommend risk taxonomy for any new risk.
- REST APIs design, development and deployment. Management of Linux environments, CI/CD Pipelines, Git Version controlling.
- Developed & deployed ETL data pipelines for a dashboard representation of business objects in Qlik.

Senior Associate Consultant (Data Analyst)
HSBC (Infosys Consultant)
01/2019 – 07/2020
London, UK

- Information Retrieval from multiple sources (Casewise Azure storage, OpenPages RESTAPIs, Cognos), Data Wrangling, Data Visualisation, Data Storytelling, Data Modelling and Data Governance.
- 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 and deployed to production.
- Managed offshore team and conducted Scrums to track sprint progress in a fast-paced Agile environment. Actively involved in Sprint planning, grooming, retrospective meetings, branching strategies, requirements gathering from stakeholders and release management for the MVP in monthly releases.

Associate Consultant
Infosys Ltd. | Client: HSBC Bank
09/2016 – 12/2018
Pune, India

- Managed 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 6.x, 7.x and 8.1 for RCM, IT Governance and ORM modules.
- Closely worked with the business stakeholders to gather stories and transform them into technical requirements in JIRA. Orchestrated the bi-weekly delivery of stories in production.

Programmer Analyst
Cognizant | Client: LPL Financials
05/2014 – 08/2016
Pune, India

- Design and development in Documentum (ECM) for the millions of documents data and metadata.
- Experience in configuration, up-gradation, and administration of Documentum Suite of Products.
- Experience in Java/J2EE, REST APIs implementation. Unit testing, & deployment of the DFC code for managing financial documents data.

🎓 EDUCATION

MSc Big Data Science
Queen Mary University of London | Grade: Distinction, 93.5%
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 App - 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, SQL

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

Big Data

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

Database

Oracle DB, SQL Server, MySQL, Amazon Dynamo DB, MongoDB, Neo4J

Cloud Computing

AWS (EC2, S3, RDS), GCP (Compute Engine, Storage, Kubernetes Engine, App Engine), Dockers, Docker-compose, Containers

Tools/IDEs

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

📋 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) Sta**r 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, Volunteering, and Interests

- **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.