Ibad Ur Rahman

DATA SCIENTIST

211 School Road, S101GN, Sheffield, United Kingdom

☐ (+44) 78-9414-6461 | ✉ ibad2762@gmail.com | 🏠 ibadia.github.io | ⓞ ibadia | 🔗 ibad2762

Summary

I have strong mathematical knowledge of Machine learning Algorithms along with proficient programming skills and good team working skills. Have worked on Deep Learning and Attention Mechanisms, have also worked on Bayesian Approach to Machine Learning. Currently working at Barclays UK as a Big Data Engineer in Elastic Stack as a Service team. Occasionally I write articles on medium explaining the concepts of Data Science in simplistic way.

Experience

Big Data Science Engineer (Elastic Stack)

BARCLAYS UK   *October. 2019 - Present*

- Working with logs coming from different sources, creating ingestion pipelines. Writing alerts on Kibana to make sure any unusual event is notified

Data Engineer

NEXDEGREE   *Jun. 2018 - September. 2018*

- Data-Pipeline and deployment of a Machine Learning Model for a micro-finance bank which gives out loan based on the User Behaviour.
- Geo-Visualisation for a bank to get the better idea of which location gives the highest return of loan.
- Natural Language Processing, Clustering Analysis for an online pharmacy to extract business insights and general user behaviour.
- Explaining the model results to the client and suggesting business benefits the organization can get using Machine Learning.

Education

The University of Sheffield   *Sheffield, United Kingdom*

MSc IN DATA ANALYTICS   *Sept. 2018 - Sept. 2019*

- Text Processing, Machine Learning and Adaptive Intelligence (in Python), Statistical Data Science in R, Industrial Team Project ,Natural Language Processing, Scalable Machine Learning, Professional Issues, Parallel Computing with GPU, Individual dissertation project
- Acheived overall Distinction in First Semester.
- Dissertation Topic: Interpretable Probabilistic Deep Learning for Text Classification which includes applying bayesian modeling on Deep Recurrent Neural Networks to make the model better explainable.

FAST(National University of Computer and Emerging Sciences)   *Karachi, Pakistan*

BACHELORS IN COMPUTER SCIENCE   *Aug. 2014 - May. 2018*

- Main courses include: Software Engineering, Data Science, Object Oriented Programming, Algorithms, Data Structures, Introduction to Operating Systems, Computer Networks, Introduction to Programming, Discrete Structures, Database Design, Information Processing Techniques

Projects

Interpretable Probabilistic Deep Learning for Text Classification

THE UNIVERSITY OF SHEFFIELD   *June. 2019 - September. 2019*

- Applied Variational Inference and MCDropout approach on Encoder Decoder Architecture in order to make the model for text classification more interpretable. Applied Bayesian Machine learning techniques on RNN along with Attention Mechanism to make it predict the sentiment of a text much better with explanation given by the model of its prediction. This can be also be applied to healthcare related machine learning as there is a dire need of explanation for a prediction done by the model. This techniques does not depends completely on the model but rather gives a measure of confidence and also highlights the parts of the input that were responsible for the particular prediction.

Validation of Scoring Criteria and Scoring Thresholds within GWEEK's Speech Intelligence

GWEEK'S SPEECH INTELLIGENCE   *Feb. 2019 - May. 2019*

- Objectives of this project were identifying the existence of any relationship between GWEEK score and Vocabulary, and classifying between Read and Planned Speech. As a result of the first task, it was found that Vocabulary did not have a major impact on the GWEEK score. For second task, Deep Learning, as well as classical Machine Learning approaches were applied on raw audio data and extracted text data from the audio. The classical machine learning methods on extracted text fetched better results than the Deep Learning approach on raw audio, which implies that extracted features given to model work better than raw features like audio.

Detailed Analysis of Bitcoins Transactions using Big-Query

Kaggle *Feb. 2018 - March. 2018*

- Detailed analysis of 821GB of bitcoin transactions data available on Kaggle through Google Big Query.
- Analysed the trends in transactions and some of its unique characteristics using plotly, Google BigQuery api, and python notebook.
- Link to analysis: `https://www.kaggle.com/ibadia/bitcoin-101-bitcoins-and-detailed-insights`.

Nex Generation Panel

Bachelors Final Year Project *Aug. 2017 - May. 2018*

- A web interface with Hadoop distributed framework on backend to analyze big data of different types.
- Can process any amount of data of different types or of any domain and give analytics.
- Can perform data segmentation using generic code for any data
- Tools Used: Python, javascript, d3.js, hadoop, HDFS

Skills

Data Science	NLP,Neural Networks, Bayesian Approach to Regression, Predictive Modeling, Clustering, Model Optimization
Deep Learning Libraries	pytorch, tensorflow, keras, scikit-learn
Data Pre-processing /Visualization Libraries	NLTK, pandas, numpy, seaborn, plotly, folium(for geospatial visualization)
Big-Data	Elastic Stack (ELK), pyspark
Programming	Python(Expert),R,C(proficient),SQL,MySQL,C#,LaTeX,Django,MVC, .NET
Cloud	Amazon Web Services,Google Colab, Google BigQuery

Publications

Artificial Neural Networks explained with code without matrices

Medium *Jan. 2018*

- Article published on Medium explaining the basics of neural networks with code. Link to article: https://medium.com/@ibaad/neural-networks-code-without-matrix-ee55e5e9bbde

Linear Regression: A Maximum Likelihood Approach

GOOD AUDIENCE *Jan. 2018*

- Article published on Good Audience Magazine through Medium explaining how Linear Regression can be viewed using a Maximum Likelihood Approach instead of the naive approach.

Bitcoin 101: How transaction works in Bitcoin

ALTCOIN MAGAZINE *Jan. 2018*

- Article published on ALTCOIN MAGAZINE through medium discussing how basic bitcoin transactions work in simplistic way for people with less technical knowledge of Blockchain.

Program Committees

2018	**Secretary**, Data Science Society, University of Sheffield
2017	**Teaching Assistant -Big Data**, NEDUET
2016	**Head**, Gaditek Dockers Challenge(DevDay 2016)
2015	**Member**, Speed Programming Competition (Procom 2015)
2014	**Member**, Speed Programming Competition (Procom 2014)

Honors & Awards

2019	**Mentioned for Industrial work at the University website**, University of Sheffield
2018	**Winner**, Speed Programming Competition,IBA Probattle
2018	**Runner Up**, Speed Programming (DEVELOPERS DAY)
2016	**Winner**, Programming Competition,(CODERS CUP) by ACM
2015	**Winner**, CBM Programming Competition
2015	**Winner**, FAST INTRAMUN (Debating)