

Eshban Suleman

eshban9492@gmail.com | +92-304-4455675

SKILLS

- Python (NumPy, Pandas, Flask)
- SQL (PostgreSQL, BigQuery)
- Git, AWS/GCP, Docker, Kubernetes
- Machine Learning (sklearn, statsmodel)
- Deep Learning (Tensorflow, Keras)
- NLP (HuggingFace, spaCy, Gensim, NLTK)
- Exploratory Data Analysis
- Predictive Modelling
- A/B Testing • Time Series Analysis
- Data Visualization (Matplotlib, Seaborn, Plotly, Bokeh, Streamlit)
- MLOps (DVC, MLFlow, Docker, GitHub Actions)

LEARNING

- Apache Airflow
- R

EDUCATION

GC UNIVERSITY

B.S. IN COMPUTER SCIENCE
2015-2019 | Lahore, Pakistan

CERTIFICATIONS

- Deep Learning Specialization
- TensorFlow in Practice Specialization
- Applied Data Science Specialization
- Data Engineering, Big Data, and Machine Learning on GCP Specialization

ARTICLES

- An Introduction to Natural Language Processing
- How to Collect, Parse, and Clean Text Data?
- Feature Engineering in NLP
- Event Driven Data Pipelines in AWS
- Introduction to Data Pipelines

LINKS

LinkedIn:// [eshban-suleman](#)
Twitter:// [@EshbanSuleman](#)
Github:// [EshbanTheLearner](#)

INTERESTS

- Book Reading • Table Tennis
- Cooking • Chess • Music • Movies

EXPERIENCE

FORETHETA | SENIOR DATA SCIENTIST

Nov 2019 – Present

- **Lead the research and development** of a semantic search engine
- **Developed tests and benchmarks** to evaluate the performance of our ML models
- **Improved the performance** of our production ML model by **49%** in one case and by **6%** in another by **improving data quality** and **hyperparameter optimization**
- **Decreased the latency** of our API by **40%** by optimizing the architecture for inference time
- **Reduced costs** of our APIs by **72%** in one case and **68%** in another by developing **automated self-replacing spot servers** in AWS
- **Brought deployment time down** by **83%** by implementing a robust **CI/CD pipeline** using **DVC, Docker, GitHub Actions** and **AWS Elastic Beanstalk**
- **Designed and developed** data pipelines to fetch data and update our online ML models
- **Took additional responsibilities** such as **interviewing, onboarding, mentoring, planning product roadmaps, delegating tasks** etc.

PRIVATE | FREELANCER

Jan 2018 – Present

Worked on following technologies and domains as a freelancer

- Big Data - Spark
- Exploratory Data Analysis
- Data Engineering w/ GCP
- Data Analysis
- Machine Learning & Deep Learning

PROJECTS

TRAINDEX

Traindex by Foretheta is AI powered semantic search engine.

- Owned Traindex at Foretheta and made design and architectural decisions
- Trained, deployed and maintained Machine Learning models
- Designed, developed and maintained data pipelines

MOVIE RECOMMENDATION ENGINE

This was a personal learning project.

- Analyzed the TMDB data and created features in form of tags (description, cast, crew)
- Used Sentence Transformer model to extract feature embeddings and cosine similarity to find similar content
- Created a simple UI in Streamlit and deployed the app to the cloud

REGINALD VON DOOM

Reginald is text summarizer for your legal and financial documents.

- Used Google's T5 model for the summarization process
- Created a simple UI in Streamlit and deployed the app to the cloud