HASSAN IFTIKHAR

Data Scientist

hassaniftikhar62@gmail.com • (16) Hassan Iftikhar | LinkedIn • Github: https://github.com/Micro046

Summary

Data Scientist with experience in machine learning, deep learning, reinforcement learning, and generative AI. Skilled at building end-to-end data solutions, from pipelines to production models, to drive measurable business impact

Experience

Inspired Square FZE
Dubai, United Arab Emirates

Data Scientist
07/2022 - 08/2024

A mobile gaming company focused on developing data-driven, engaging puzzle and strategy games.

- Designed and optimized ETL pipelines to source and structure data in cloud databases.
- Analyzed large datasets to extract actionable insights supporting product and business strategy.
- Developed and deployed machine learning models for predictive analytics and game performance optimization.
- Built interactive dashboards and visualizations to monitor KPIs and guide decision-making.
- Conducted data-driven analyzes to increase revenue and improve player engagement metrics.
- Integrated new data pipelines with existing databases, streamlining data sourcing and accessibility.
- Executed A/B testing and user segmentation to deliver targeted experiences and boost key KPIs.

Applied AI Center, Skoltech
Moscow, Russia

Researcher
09/2024 - Present

A research hub at Skolkovo Institute driving innovation in AI and machine learning.

- Conducted research on Umbrella Reinforcement Learning (URL), developing an offline RL variant to extend its applicability.
- Designed and experimented with RL algorithms for efficient policy learning.
- Contributed to the theoretical understanding and practical adaptation of advanced RL and generative AI techniques.

Sberbank
Moscow, Russia

Data Scientist
05/2025 - 08/2025

Sber AI — Sberbank's AI division, building generative models and intelligent systems for finance and beyond.

- Designed and developed the speaker feedback system, enabling real-time verbal delivery insights.
- Built audio-processing pipelines for feature extraction (pitch, pace, clarity) and preprocessed speech data.
- Trained ML models to assess speech quality, refine performance, and provide immediate feedback on delivery attributes.
- Integrated feedback mechanisms to guide users on pacing, tone, and clarity in live speaker scenarios.
- Deployed the system via an interface or API, ensuring accessibility for coaching or presentation tools.

DCUBE Tech
Islamabad, Pakistan

Data Engineer
01/2022 - 05/2022

Pakistan-based firm specializing in computer vision, AI, and OCR solutions

- Moving raw data (360 shots) in the form of batch from one server to another

Education

Skolkovo Institute of Science and Technology
Moscow

Masters in Data Science
09/2024 - 06/2026

Namal University
Mianwali

Bachelor of Science in Electrical Engineering
09/2018 - 06/2022

Skills

Machine Learning · AI · Supervised Learning · Unsupervised Learning · Reinforcement Learning · Generative AI · Agentic AI · Deep Learning · CNNs · LSTMs · UNet · Transformers · Programming · Python · NumPy · Pandas · Scikit-Learn · TensorFlow · PyTorch · SciPy · Keras · OpenCV · Matplotlib · Seaborn · MATLAB · C++ · Statistics · A/B Testing · Time Series · PCA · Bayesian Methods · Data Engineering · ETL · Talend · MySQL · Google BigQuery · MySQL Workbench · Visualization · Tableau · Google Data Studio · React · Node.js · Cloud · DevOps · GCP · Vertex AI · Cloud Run · AWS SageMaker · Docker · Tools · Git · Jupyter · Cron Jobs · VAE · Diffusion Models · GANs · Flows · SAC · PPO · Vision Transormers

Projects

Speaker Feedback System

Created a machine learning system to analyze **presentation video**s by slide, extracting features (**speech pace, clarity, tone, timing, gaze, emotions, clothing**) for speaker feedback. Built the pipeline: feature extraction, model training, real-time evaluation.

- Created AI tool for slide feedback using speech, vision, language models; **transformer-based OCR** in English/Russian. Evaluated on presentation videos, offering feedback on clarity, gaze, emotions, gestures, attire, slide content.

Retinal Vessel Segmentation and Diabetic Retinopathy detection

Developed a **UNET**-based deep learning model to segment **retinal blood vessels**, employing separate training for thick/thin vessels and a **fusion mechanism.** Enhanced thin vessel detection with optimized preprocessing, skeletonization, and vessel-width separation, outperforming baseline models.

- Achieved **96.42% accuracy, 97.13% sensitivity,** and 84.67% specificity on the DRIVE dataset, outperforming UNET, AG-UNET, and Dense Net baselines. Demonstrated improved segmentation of thin vessels a critical factor for early diabetic retinopathy diagnosis.

A/B Testing & User Segmentation for Mobile Games

Performed **user segmentation** on mobile puzzle games (***X2* and *2248***) *using* **RFM analysis** and the **20/80 rule** to identify high-value players. Designed and executed A/B tests on game economy parameters (e.g., coin rewards per level) to optimize player engagement, in-app purchases, and ad interactions.

- Increased game revenue by **2–3%**, while maintaining player **retention rates** demonstrating effective balance between monetization and user experience.

ETL Pipeline Development for Mobile Analytics

Built automated ETL pipelines to extract user and monetization data from APIs such as Apps flyer and App Lovin, ingest it into **Google Cloud Storage (buckets)**, and load it into **Big Query** for analytics. Used **Talend Data Studio**, **Pub/Sub**, and **Dataflow** to orchestrate, transform, and scale data pipelines.

- Delivered a **cloud-based data** pipeline that cut manual reporting, enhanced analyst data access, and enabled real-time user metrics insights.

User Lifetime Prediction & Segmentation

Developed a predictive modelling framework that uses a player's **first 7 days of activity** to forecast their **90-day lifetime value and category**. Designed strategies to target **high-value players** with personalized offers and promotions to maximize retention and in-app purchases.

- Enabled early identification of **golden users**, allowing the marketing team to deploy targeted offers and campaigns. Improved personalization contributed to higher retention and revenue potential, optimizing user acquisition ROI.

Certification

Coursera—Machine Learning

DeepLearning.AI — Natural Language Processing with Classification and Vector Spaces

Google — Data Analytics Certificate, Data Analytics Specialization, Finance Data Analytics Professional Certificate

KPMG — Data Analytics Consulting Virtual Internship

GE — Virtual Learning Program (Analytics)

Coursera — Diabetic Retinopathy Detection with Artificial Intelligence

Sololearn — Python for Data Science, C++