

EMRE OKCULAR

San Francisco, CA | 415 799 6601

emreokcular@gmail.com | okcular.com.tr | github.com/emreokcular | linkedin.com/in/emreokcular

Experienced senior machine learning engineer consultant with 6+ years of experience designing, building, testing, deploying and maintaining ML systems, including high-performance real-time streaming analytics and marketing applications, with a software and data engineering background. Specialized in building end-to-end scalable machine learning pipelines and MLOps. Passionate client-facing ML solution architect with strong communication skills. Solving real-world problems from a wide variety of business domains by combining domain knowledge, data science and engineering perspective.

WORK EXPERIENCE

Capgemini, San Francisco, CA

Oct, 2021 – Present

Senior Machine Learning Engineer Consultant

- Led data scientists and engineers in building analytics and machine learning applications using Python Data Science stack to support pre-sales activities such as workshops, proposals and demos for wide variety of client use cases.
 - Created a hybrid demand prediction model combining clustering and curve-fitting models trained on billions of records using PySpark on Azure Databricks and improved store planning capabilities by optimizing shelf space of departments in full fleet.
 - Built an end-to-end production-level ML pipeline including auto feature selection, feature engineering, preprocessing, and evaluation stages to optimize store planning and increased annual sales by 10%.
 - Crafted and optimized steps of the deployed ML pipeline that processes high volume delta tables with PySpark by code-level and cluster-level optimization techniques and reduced execution time by 20%.
 - Deployed an MLOps process using MLFlow by registering the model as endpoints and serving the model as a REST API.
 - Created neural networks to diagnose river blindness disease from biopsy images using PyTorch on AWS Sagemaker.
- Increased model accuracy by fine tuning pre-trained state-of-art object detection models and data augmentation techniques.

Dictionary.com, Oakland, CA

Data Scientist

Jan, 2021 – Aug, 2021

- Predicted click-through rate with random forest classifier trained on website logs and cookies. Increased model accuracy by 5% using sampling and feature engineering techniques in scikit-learn pipelines within AWS (S3, EC2, EMR, RDS, Athena).
- Identified most engaged user segments by exploring website behavior to gain insights for improving the ad auction system.

Turkcell, Istanbul, TURKEY

Data Engineer

Aug, 2019 – Oct, 2020

- Achieved the ability to analyze streaming data in real-time and take immediate actions through communication channels by developing big data processing systems with Java, Python and Agile Methodologies.
- Expanded event-based scenarios such as gamification, anti-churn, up-sell, and retention resulted in a 15% increase in annual revenue by integrating big data sources using Kafka, Spark, Logstash and SQL into complex event processing systems.
- Increased monthly bundle package sales by 10% discovering customers' opinions from messages in SMS channels with NLP techniques such as named entity recognition, sentiment analysis, and text classification using PyTorch NN.

Software Engineer

Aug, 2017 – Aug, 2019

- Increased the daily capacity of sending millions of messages and notifications by 50% building highly scalable campaign management applications with Java, PL/SQL, and Python using best practices for software development lifecycle.
- Automated generation of daily campaign reports for all channels using a vast amount of relational data with tuned SQL queries and PL/SQL objects such as procedures, triggers, views, and tables.
- Empowered the marketing team to derive strategic insights for campaigns by creating a performant Java REST service for collecting push notification responses in the Oracle SQL database.

EDUCATION

University of San Francisco, San Francisco, CA

Master of Science in Data Science

Aug, 2020 – Aug, 2021

- Coursework: Python, Data Structures and Algorithms, Statistics, Data Visualization, SQL, NoSQL, Machine Learning, Regression, Deep Learning, NLP, Time Series, Design of Experiments(A/B Testing), Distributed Computing (Spark)

Yildiz Technical University, Istanbul, TURKEY

Bachelor of Science in Computational and Applied Mathematics, Honors Student

Sep, 2012 – Jun, 2017

Graz University of Technology, Graz, AUSTRIA

Erasmus Exchange Program in Computer Science and Mathematics, Honors Student

Oct, 2014 – Aug, 2015

TECHNICAL SKILLS

Python(Pandas, Scikit-learn, NumPy, SciPy), PyTorch, Java, SQL, NoSQL, Spark, Databricks, Kafka, Linux, Git, AWS, Azure