EMRE OKCULAR

San Francisco, CA | 415 799 6601 | emreokcular@gmail.com | okcular.com.tr | github.com/emreokcular | linkedin.com/in/emreokcular

EDUCATION

University of San Francisco San Francisco, CA
Master of Science in Data Science August, 2020 – August, 2021(expected)

- Relevant Coursework: Python, Data Structures and Algorithms, Probability, Statistics, Data Visualization, Linear Regression, Machine Learning, Deep Learning, NLP, Product Analytics, Design of Experiments, Time Series, Distributed Computing(Spark), SQL, Business Strategy, Data Ethics

Yildiz Technical University Istanbul, TURKEY
Bachelor of Engineering in Computational and Applied Mathematics September, 2012 – June, 2017

- Relevant Coursework: Linear Algebra, Calculus, Discrete Mathematics, Abstract Mathematics, Algorithms, Programming, OOP, Probability, Statistics, Data Management, Optimization, Cryptography, Computer Vision, Image Processing, AI

Graz University of Technology Graz, AUSTRIA
Erasmus Exchange Program in Computer Science and Mathematics October, 2014 – August, 2015

- Ranked 1st among the outgoing Erasmus Exchange Program students in the department and awarded EU Scholarship.

WORK EXPERIENCE

Data Science Intern January, 2021 – Present
Dictionary.com Oakland, CA

- Predicting click-through rate (CTR) with random forest classifier trained on website logs and cookies. Increased model accuracy by 5% using feature engineering techniques.
- Identified most engaged users by exploring user website behavior to gain insights for improving the ad auction process.

Software Engineer August, 2017 – October, 2020
Turkcell Istanbul, TURKEY

- Built highly scalable campaign management applications with Java, PL/SQL, and Python for the direct marketing team to be capable of sending millions of messages and notifications per day.
- Developed real-time streaming analytics systems with Lean-Agile Methodologies. Gained the ability to analyze streaming data in real-time and to take immediate actions with outbound communication channels.
- Increased monthly bundle package sales by 15% discovering customer's opinions with NLP techniques such as named entity recognition, sentiment analysis and text classification in SMS channels.
- Created a high-performance synchronous Java REST service for collecting and storing push notification responses in the SQL database to derive insightful metrics for the marketing team.
- Enabled event-based scenarios such as gamification, anti-churn, up-sell, and retention for business teams by integrating data sources to complex event processing systems and building daily ETL jobs for a vast amount of relational data.

Data Science Intern February, 2017– May, 2017
EVAM Streaming Analytics Istanbul, TURKEY

- Researched streaming ML algorithm papers and identified algorithms to implement after understanding business needs.
- Enabled real-time segmenting ability, including outlier detection for client use cases by implementing density-based streaming clustering algorithm DenStream to the product with Java.
- Performed tests with data from different clients in R&D processes with R, Python, Java, Apache Spark, and MOA.

DevOps Intern June, 2016 – July, 2016
Anadolu Insurance Istanbul, TURKEY

- Collaborated with the DevOps team to develop automated pipelines with Jenkins for application lifecycle management.

PROJECTS

- **Cancer Classification by Liquid Biopsy**
  Achieved 73% average accuracy with different models and placed in the top 10 in Kaggle private leaderboard. Fit and tuned various scikit-learn classifiers such as boosting and PyTorch deep neural networks to predict multi-class cancer types by identifying top features from microRNA profiles.
- **ML algorithm implementations from scratch in Python**
  Regularized Linear and Logistic Regression with Gradient Descent, Naive Bayes, Decision Trees, Random Forest, K-means Clustering, Boosting, Deep Neural Networks, recommendation engine, and automated feature selection algorithms.
- **Ad Click Prediction**
  Predicted ad clicks from KDD Cup dataset achieving 83% accuracy. Various classifier models, including tree-based ensemble methods used in scikit-learn pipelines applying feature selection, preprocessing, and tuning techniques with CV.

TECHNICAL SKILLS

Python (Pandas, Scikit-learn, NumPy), PyTorch, Java, R, PL/SQL, MongoDB, Spark, Kafka, Linux, Git, Docker, AWS, Jenkins