

# EMRE OKCULAR

San Francisco, CA | 415 799 6601 | [emreokcular@gmail.com](mailto:emreokcular@gmail.com) | [okcular.com.tr](http://okcular.com.tr) | [github.com/emreokcular](https://github.com/emreokcular) | [linkedin.com/in/emreokcular](https://linkedin.com/in/emreokcular)

## EDUCATION

<b>University of San Francisco</b> Master of Science in Data Science	San Francisco, CA August, 2020 – August, 2021(expected)
<ul style="list-style-type: none"><li>Relevant Coursework: Python, Data Structures and Algorithms, Probability, Statistics, Data Visualization, Machine Learning, Deep Learning, NLP, Product Analytics, Design of Experiments, Time Series, Distributed Computing(Spark), SQL, Business Communication, Data Ethics</li></ul>	
<b>Yildiz Technical University</b> <i>Bachelor of Engineering in Computational and Applied Mathematics</i>	Istanbul, TURKEY September, 2012 – June, 2017
<ul style="list-style-type: none"><li>Relevant Coursework: Linear Algebra, Calculus, Discrete Mathematics, Abstract Mathematics, Algorithms, Programming, OOP, Probability, Statistics, Data Management and File Structures, Optimization, Cryptography, Operations Research, Computer Vision, Image Processing, AI</li></ul>	
<b>Graz University of Technology</b> Erasmus Exchange Program in Computer Science and Mathematics	Graz, AUSTRIA October, 2014 – August, 2015
<ul style="list-style-type: none"><li>Ranked 1st among the outgoing Erasmus Exchange Program students in the department and awarded EU Scholarship.</li></ul>	

## WORK EXPERIENCE

<b>Data Science Intern</b> <i>Dictionary.com</i>	January, 2021 – Present Oakland, CA
<ul style="list-style-type: none"><li>Predicting click-through rate (CTR) with random forest classifier trained on website logs and cookies. Increased model accuracy by 5% using feature engineering techniques.</li><li>Identified most engaged users by exploring user website behavior to gain insights for improving the ad auction process.</li></ul>	
<b>Software Engineer</b> <i>Turkcell</i>	August, 2017 – October, 2020 Istanbul, TURKEY
<ul style="list-style-type: none"><li>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 per day.</li><li>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.</li><li>Increased monthly bundle package sales by 15% discovering customer's opinions with text classification techniques in SMS channels.</li><li>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.</li><li>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.</li></ul>	
<b>Data Science Intern</b> <i>EVAM Streaming Analytics</i>	February, 2017– May, 2017 Istanbul, TURKEY
<ul style="list-style-type: none"><li>Explored streaming ML algorithm papers and identified algorithms to implement after understanding business needs</li><li>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.</li><li>Performed tests with data from different clients as R&amp;D processes with R, Python, Java, Apache Spark, and MOA.</li></ul>	
<b>DevOps Intern</b> <i>Anadolu Insurance</i>	June, 2016 – July, 2016 Istanbul, TURKEY
<ul style="list-style-type: none"><li>Collaborated with the DevOps team to develop automated pipelines with Jenkins for application lifecycle management.</li></ul>	

## PROJECTS

- Cancer Classification by Liquid Biopsy**  
Achieved 73% average accuracy with 21 different models and placed in the top 10 leaderboard in Kaggle. Fit various scikit-learn classifiers and PyTorch neural networks to predict multi-class cancer types by identifying top features.
- ML algorithm implementations from scratch in Python**  
Regularized Linear and Logistic Regression, Naive Bayes, Decision Trees, Random Forest, K-means Clustering, Gradient Boosting, Deep Neural Networks, recommendation engine, and automated feature selection algorithms.
- Ad Click Prediction**  
Predicted ad clicks from KDD dataset with different classifier models in scikit-learn pipelines. Achieved 83% accuracy applying feature selection, preprocessing, and tuning techniques.

## TECHNICAL SKILLS

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