

Emin Can Oguz

Istanbul/Turkey
emincanoguz11@gmail.com
emincanoguz11.github.io

ABOUT ME

As a data engineer with 3 years of experience, I specialize in designing, building, and maintaining high-performance data systems. I have experience managing large-scale data pipelines, implementing streaming solutions, and optimizing costs and queries within clusters. In particular, I have successfully managed data infrastructure for multiple Black Friday events, delivering reliable performance and insights to stakeholders.

EDUCATION

Maltepe University- Software Engineering (2016-2021)

WORK EXPERIENCE

Hepsiburada - Data Engineer (April 2021 - Present)

- I developed a Data Mesh pipeline for all teams, as well as a Data pipeline. Additionally, I implemented a streaming job utilizing real-time technologies, and managed a cluster exceeding 10PB. To further support my teams, I also created Airflow pipelines. I focused on optimizing costs and queries within the cluster. To obtain a data engineering position with a focus on optimizing data pipelines and infrastructure for high-traffic events, such as Black Friday.

Doctorla aka Online Saglikcim - Junior Data Scientist (April 2020 -Jan 2021)

- As a junior data scientist with a background in machine learning, I have experience developing and deploying NLP models using Tensorflow. I am passionate about using my skills to solve real-world problems, particularly in the healthcare domain. I have experience working with large-scale datasets and utilizing advanced techniques to extract insights and create predictive models.

RemoteTeam.com - Junior Data Scientist (May 2020 - Jan 2021)

- As a Junior data scientist with a background in recommendation systems, I have experience deploying algorithms to generate personalized recommendations for users. I have experience working with large-scale datasets and utilizing advanced techniques to extract insights and create predictive models.

Yesil Science - Intern (April 2020 – June 2020)

- I developed Mandarin project for scoring healthcare using Machine Learning.

SKILLS

Python, Java, ,Scala, C#, SQL, NoSQL, Cloud Technologies (AWS,Google Cloud), Prometheus, Apache Druid, Apache Spark, Apache Hadoop, Apache Kafka, RabbitMQ, Linux, Kubernetes, Docker, Gitlab CI/CD, Jenkins, Tensorflow, Spark NLP, Scikit-Learn, Google BigQuery, Google Cloud Storage. Apache Airflow, DBT, Amundsen. Datahub, MongoDB