Achhuta Nand Jha

Data Scientist

📱 +90 7281031975

🌐 India-Mathura

🅔 anjha1

✉ achhuta.jha$_{cs.da23@gla.ac.in}$

in Achhuta Nand Jha

⯁ anjha1

About me

I am a T-shaped Data Scientist who has general knowledge and experience on different IT areas such as Machine Learning Engineering, Amazon Web Services, System Engineering, Back-End Development, Data Engineering and Unit Testing. I have more than 6 years of experience in Python.

My passion is to bring innovation and value through cross-functional collaboration. I follow Design Principles to build real-world Data Science projects with end-to-end pipelines. I am also working on integrating MLOps practices into my work. Beyond these aspects, I am not only an IT Engineer but also an eager lifelong learner.

Besides my technical expertise, I am an extroverted and open-minded person with excellent communication skills. I leverage both my soft and hard skills to position myself as a leading engineer in the industry.

Experience

Dec 2021 Present	**Data Scientist**	**Hepsiemlak**

- Predicting sales and rental prices of real estates using Machine Learning. Finding low-priced real estates. Building up a REST API that predicts prices of real estates via ML models.
  Stack: Python, LightGBM, PyTest, FastAPI, Hadoop, AWS, Docker.
- Creating several image regression, segmentation & classification projects on image data.
  Stack: Python, PyTorch, Transfer Learning, RabbitMQ, Docker, AWS.
- Building up a Fraud Detection System for Affiliate Marketing project.
- Computing price indexes of locations using Apache Spark.

Oct 2019 Dec 2021	**Data Scientist**	**Urbanstat**

- Created an ML based risk scores of automobiles & houses for the use of insurance companies. Lowered loss ratios up to 7%.
  Stack: Python, Flask, LightGBM, PostgreSQL, Docker.
- Predicted wildfires and dangerous regions in California & West Coast of USA using ML algorithms. Outperformed predictions of state government. Used SHAP library for interpretable outputs.
- Developed an ML based churn analysis for house policies.

Sep 2018 Oct 2019	**Fraud Systems Engineer**	**Turkcell**

- Took part in an ML project that aimed to predict which customers were going to object to their invoices. Dealt with an unbalanced problem.
- Monitored Fraud & Credit Control services of Turkcell. Managed several Physical & Virtual Machines. Deployed software packages on live environment. Used Linux and SQL a lot.

Apr 2018 Aug 2018	**Junior Data Scientist**	**Organon Analytics**

- Designed dashboards for end users via Shiny library of R.
- Dealt with Customer Segmentation Analysis through Clustering.

Education

2012 Sep 2018 Jan	**Bachelor of Science, Industrial Engineering**	**Boğaziçi University**

- Took Statistics, Data Mining & Time Series Forecasting courses.
- Forecasted hourly electricity consumption of a specific region using different Machine Learning Algorithms in my graduation project.

Achievements

Kaggle	**Kaggle Expert**

- Collected 3 silver and 2 bronze medals. Joining 20+ competitions.
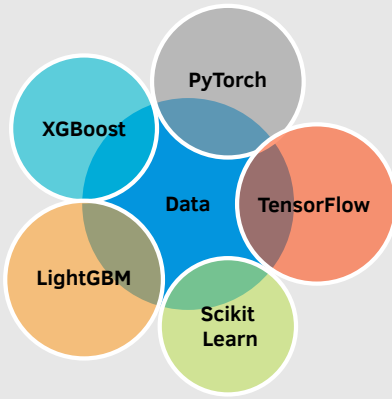
2017 ALES	**Ranked 147th in Turkey**

- Achieved a ranking of 147th out of 300 thousand participants in the 2017 Spring ALES (Academic Personnel and Postgraduate Education Entrance Exam).

2012 YGS	**Ranked 300th in Turkey**

- Attained a position of 300th among 1.86 million candidates in the 2012 YGS (The Transition to Higher Education Examination).

ML Libraries

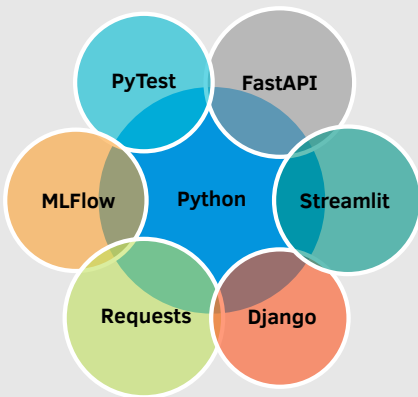## Programming Skills

- Python • Java

- HTML • CSS • JavaScript

- Linux • SQL • Docker

- Microsoft Office • MLOps • Git

- AWS • Airflow • Kubernetes

Python Libraries

Technical skills

- Data Analytics, Data Visualization

- Machine Learning

- Database Management System

- IBM Cloud (Watson Studio)

- Big Data Tools (e.g., Spark, Hadoop)

Projects

Django App	**muhammedbuyukkinaci.com**	**GitHub Link**

- A Django Application deployed on a VPS of DigitalOcean.

Time Series	**Bitcoin Trading**	**GitHub Link**

- Built up an LSTM model to predict price volatility on 4-hourly Bitcoin data.
- Obtained a fractional signal for trading.

Image Segmentation	**DeText**	**GitHub Link**

- Created a simulated dataset by overlaying text on images.
- Trained a UNet model to remove text from images via TensorFlow.

Certificates

AWS	**Introduction to Amazon Web Services**	**Certificate Link**

- Completed a 25-hours AWS Course on Udemy.
- Learned fundamentals of AWS and services like EC2, S3, IAM, VPC, RDS, ECR, EKS and DynamoDB etc.

Docker	**Docker A-Z™**	**Certificate Link**

- Completed a 16-hours course on Udemy. Learned basics of Docker.

Kubernetes	**Kubernetes Basics**	**Certificate Link**

- Completed a 17-hours Kubernetes Course on Udemy.
- Learned fundamentals of Kubernetes and K8s objects such as Pod, Deployment, Service, Ingress, Cronjob etc.

MLOps	**Complete MLOPS Bootcamp**	**Certificate Link**

- Completed a 3.5 hours MLOps course.
- Learned useful tools for MLOps practices.

Linux	**Linux A-Z™**	**Certificate Link**

- Completed a 10-hours Linux Course on Udemy.
- Elevated my Linux knowledge to an advanced level.

Big Data	**Big Data A-Z™**	**Certificate Link**

- Completed a 12-hours Big Data Course on Udemy.
- Learned basics of Apache Kafka, Apache Hive, Hadoop, Apache Spark and other NoSQL technologies.

Airflow	**Introduction to Apache Airflow**	**Certificate Link**

- Completed a 2.5 hours Airflow course.
- Learned fundamentals of Airflow.