THE DATA
SCIENCE BOOTCAMP

Mastering Data Science With Deep Learning

Syllabus

Instructed by Zafer Acar
zacar@mtu.edu

COURSE DESCRIPTION

The Mastering Applied Data Science with Deep Learning course is designed to provide students with a holistic learning experience that encompasses the key pillars of data science, project-based learning, and deep learning.

In the Data Science program, you will not only learn the fundamental concepts and techniques of data science but also apply them in real-world scenarios through a series of hands-on projects and assignments. This practical approach will solidify your understanding and equip you with the necessary skills to tackle complex data challenges.

Project Based Learning takes your knowledge to the next level by guiding you through the process of developing end-to-end machine-learning projects from scratch. You will gain invaluable experience in problem-solving, data preprocessing, model selection, and evaluation as well as learn how to communicate your findings effectively.

The Deep Learning component of the course delves into the realm of neural networks and explores advanced deep learning architectures. You will learn how to design and train deep neural networks, optimize model performance, and leverage cutting-edge techniques such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs).

By combining these three modules, you will not only acquire a solid foundation in data science but also gain practical project experience and develop a deep understanding of the principles and applications of deep learning. This comprehensive skill set will empower you to tackle complex data science problems and excel in the field of applied data science.

In addition to the comprehensive coverage of data science, project-based learning, and deep learning, the Mastering Applied Data Science with Deep Learning course also introduces the essential concept of MLOps that focuses on the operationalization and deployment of machine learning models at scale.

Seize AI Opportunities in the $15.7 Trillion Market Learn Real World Use Case

Data Science, Machine Learning, AI Bootcamp

Data science has emerged as one of the most promising professions for success in the 21st century. Our immersive Data Science Bootcamp is meticulously crafted to equip you with the essential skills required to embark on a thriving career or project in data science and artificial intelligence. Our comprehensive curriculum provides you with a strong foundation and invaluable insights, which are further reinforced through 60 practical, hands-on in-class projects. Additionally, we offer the opportunity to work on real-life business projects, diving deep into the practical applications of data science.

Basics

- Statistics 101
- Python 101
- Introduction to Data Science with Python
- Exploratory Data Analysis
- Data Visualization & Information Analysis
- Data Munging
- Feature Engineering

Machine Learning

- Introduction to Machine Learning
- Supervised Learning, Unsupervised Learning, Reinforcement Learning
- Linear, Polynomial, Ridge, Lasso Regression, ElasticNet, Extra Tree Regressors
- Intermediate Machine Learning
- Classification and clustering
- Naive Bayes, Bernoulli, Decision Tree Classifiers
- Advanced Machine Learning, Anomaly Detection
- Metrics to evaluate predictions
- Support Vector Machines
- Principal Component Analysis
- Train Test Split, K-Folds, Accuracy, Recall, Confusion Matrix

Artificial Intelligence (AI)

- NLP - Natural Language Processing and sentiment analysis
- Big Data with Hadoop and Spark
- Recommendation Systems

  - Collaborative Filtering
  - Content-Based Filtering
  - Hybrid Recommender Systems
  - Matrix Factorization
  - Deep Learning in Recommender Systems
  - Cold Start Problem
  - Real-World Applications

- Computer Vision with OpenCV - Face Recognition
- Deep Learning, Tensorflow - Time Series
- Artificial Neural Networks (ANN)
- Deep neural networks (DNN)

- Convolutional Neural Networks (CNN)
- Popular CNN Architectures: LeNet-5, AlexNet, VGGNet, GoogLeNet,
- ResNet Recurrent Neural Networks (RNN)
- Generative Adversarial Networks (GAN) for generative modeling
- Long Short-Term Memory (LSTM)
- 3D computer vision and depth estimation
- Social network analysis, page ranking
- Graph Theory
- Automated Computer Vision

Automated Machine Learning (AutoML)

■ **Pycaret:** PyCaret simplifies the machine learning workflow and accelerates the model development process by automating several steps

■ **AutoKeras:** It automates the time-consuming process of manual model selection and hyperparameter tuning

■ **H2O.ai:** Offers AutoML capabilities through its H2O AutoML component

■ **Dask-ML:** Integrates with Dask to provide distributed computing capabilities

Reinforcement Learning(RL)

Introduction to Reinforcement Learning:

■ Concepts and Algorithms
■ Applications and Challenges
■ Exploration vs. Exploitation
■ Policy Gradient Methods
■ Sequential Decision Making
■ Multi-agent Reinforcement Learning

Geographic Information Systems (GIS)

■ Introduction to Geographic Information Systems (GIS)
■ Spatial Data Analysis
■ Geospatial Data Visualization
■ Web GIS and Cloud-Based Mapping
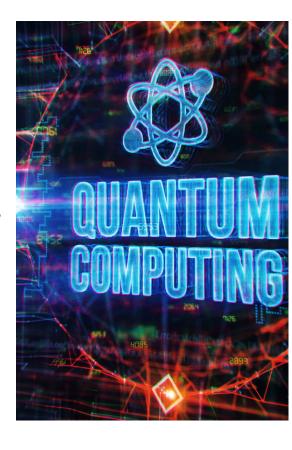■ Remote Sensing and GIS Integration

Quantum Machine Learning (QML)

- Quantum Computing
- QML algorithms
- Qiskit Aqua
- IBM Quantum Experience

Transfer Learning and Transformers

- Leveraging Pretrained Models for Faster and More Accurate Machine Learning Inductive transfer learning
- Transductive transfer learning
- Unsupervised transfer learning
- Common Applications of Transfer Learning
- Best Practices for Transfer Learning
- Transformers and their role in NLP
- Self-attention mechanism
- Transformers vs. Recurrent Neural Networks (RNNs)

Large Language Models (LLMs)

- A Deep Dive into Large Language Models
- Language Generation at Scale
- Prompt Engineering
- Next-Level Natural Language Processing
- Creative Text Generation with Large Language Models

Machine Learning Operations (MLOps)

- Bridging the Gap Between Machine Learning and Operations
- Building End-to-End ML Pipelines
- Continuous Integration and Continuous Deployment (CI/CD) for Machine Learning
- Data Governance and Compliance in MLOps
- Version control, Containers, Docker and Kubernates
- Building Intelligent Mobile Apps
- Cross-Platform ML Model Deployment

Why should you take this class?

In-person Training

Applied Labs; an Innovative Way to Learn

100% Hands-On Learning

Mentoring, One-on-One Project Based Learning

Certificate after completion

_____

Instructed by Zafer Acar

Module	Topics	Hours	Platform
Week 1	• Statistics 101 • Python 101 • Introduction to Data Science with Python • Exploratory Data Analysis • Data Visualization & Information Analysis • Data Munging • Feature Engineering	14	**Anaconda Jupyter Notebooks**
Week 2	• Machine Learning • Supervised Machine Learning • Unsupervised Machine Learning • Reinforcement Learning	14	**Anaconda Jupyter Notebooks**
Week 3	• NLP - Natural Language Processing and sentiment analysis	14	**Anaconda Jupyter Notebooks & Huggingface.co**
Week 4	• Computer Vision with OpenCV - Face Recognition • Deep Learning, Artificial Neural Networks (ANN) •	14	**Anaconda Jupyter Notebooks**
Week 5	• Recommendation Systems • Big Data Hadoop Spark	14	**Google Colab - Apache Spark**

Module	Topics	Hours	Platform
Week 6	• Natural Language Processing	14	**Anaconda Jupyter Notebooks**
Week 7	• Automated Machine • Learning (AutoML) Geographic Information Systems • Quantum Computers and Machine Learning (QML)	14	**Google Colab IBM Qiskit**
Week 8	• Transfer Learning and Transformers • Large Language Models (LLMs) • Reinforcement Learning(RL) • Machine Learning Operations (MLOps)	14	**Visual Studio Code Pythonanywhere Github Huggingface.co**
Week 9	• Project Based Learning 1 (PBL1) • Project Solutions and Presentations by the Students	2	**Zoom**
Week 10	• Project Based Learning 2 (PBL2) • Project Solutions and Presentations by the Students	2	**Zoom**
Week 11	• Project Based Learning 3 (PBL3) • Project Solutions and Presentations by the Students	2	**Zoom**
Week 12	• Project Based Learning 4 (PBL4) • Project Solutions and Presentations by the Students	2	**Zoom**
Week 13	• Project Based Learning 5 (PBL5) • Project Solutions and Presentations by the Students	2	**Zoom**
Week 14	• Independent Studies	2	**Zoom**
Week 15	• Independent Studies	2	**Zoom**
Week 16	• Capstone project presentations (families and friends will join)	2	**Zoom**

In our Data Science Bootcamp, we will be utilizing the following websites, and it is necessary for all attendees to create accounts on these platforms:

1. **GitHub:** Please sign up at www.github.com to create an account.
2. **Kaggle:** To participate in Kaggle competitions and access data sets, please sign up at www.kaggle.com.
3. **IBM Quantum Platform:** To work with quantum computing resources, please sign in at https://quantum-computing.ibm.com/ and create an account.
4. **MIT App Inventor:** To create mobile apps, please sign up at https://appinventor.mit.edu/ and create an account.
5. **Google Colab:** To collaborate on and run Python notebooks in the cloud, please sign in at https://colab.research.google.com/ using your Google account.
6. **Amazon Web Services(AWS) Machine Learning:** To utilize AWS machine learning services, please sign up at https://aws.amazon.com/machine-learning/ and create an account.
7. **PythonAnywhere:** To host, run, and code Python in the cloud, please sign up at https://www.pythonanywhere.com/ and create an account.
8. **Hugging Face:** To access pre-trained models and collaborate with the AI community, please sign up at https://huggingface.co/ and create an account.

Please ensure that you have created accounts on these platforms prior to the start of the boot camp to make the most of the learning experience.