

# ENG 346

# Data Structures and Algorithms for Artificial Intelligence

## Course Overview

Dr. Mehmet PEKMEZCI

[mpekmezci@gtu.edu.tr](mailto:mpekmezci@gtu.edu.tr)

<https://github.com/mehmetpekmezci/GTU-ENG-346>

## Quick Notes : PROGRAM – ALGORITHM – MATH RELATION

- A Computer Program is based on an Algorithm.
- Algorithm (Webster Dict.) :

An algorithm is a finite sequence of **mathematically** rigorous instructions, typically used to solve a class of specific problems or to perform a computation.

# Quick Notes : MATH. > ALGO. > PROG. LANGUAGE

- Example : Fibonacci Numbers :  $a_n = a_{n-1} + a_{n-2}$
- $a_{100000} = ?$
- **Math** : Binet's Formula (Generating Functions)  $O(\log(n))$

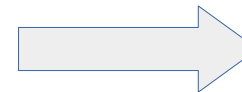
$$F(n) = \frac{\phi^n - (1-\phi)^n}{\sqrt{5}}$$

$$\phi = \frac{1 + \sqrt{5}}{2}$$

- **Algorithm** : Find an algorithm that calculates faster with less resource :

```
def nth_fibonacci(n):  
    if n <= 1: return n  
    return nth_fibonacci(n - 1) + nth_fibonacci(n - 2)  
  
print(nth_fibonacci(5))
```

$O(2^n)$



```
F_n=0 ; F_n_1=2; F_n_2=1  
n=5  
for in range(n):  
    F_n = F_n_1 + F_n_2  
    F_n_2=F_n_1  
    F_n_1=F_n  
print(F_n)
```

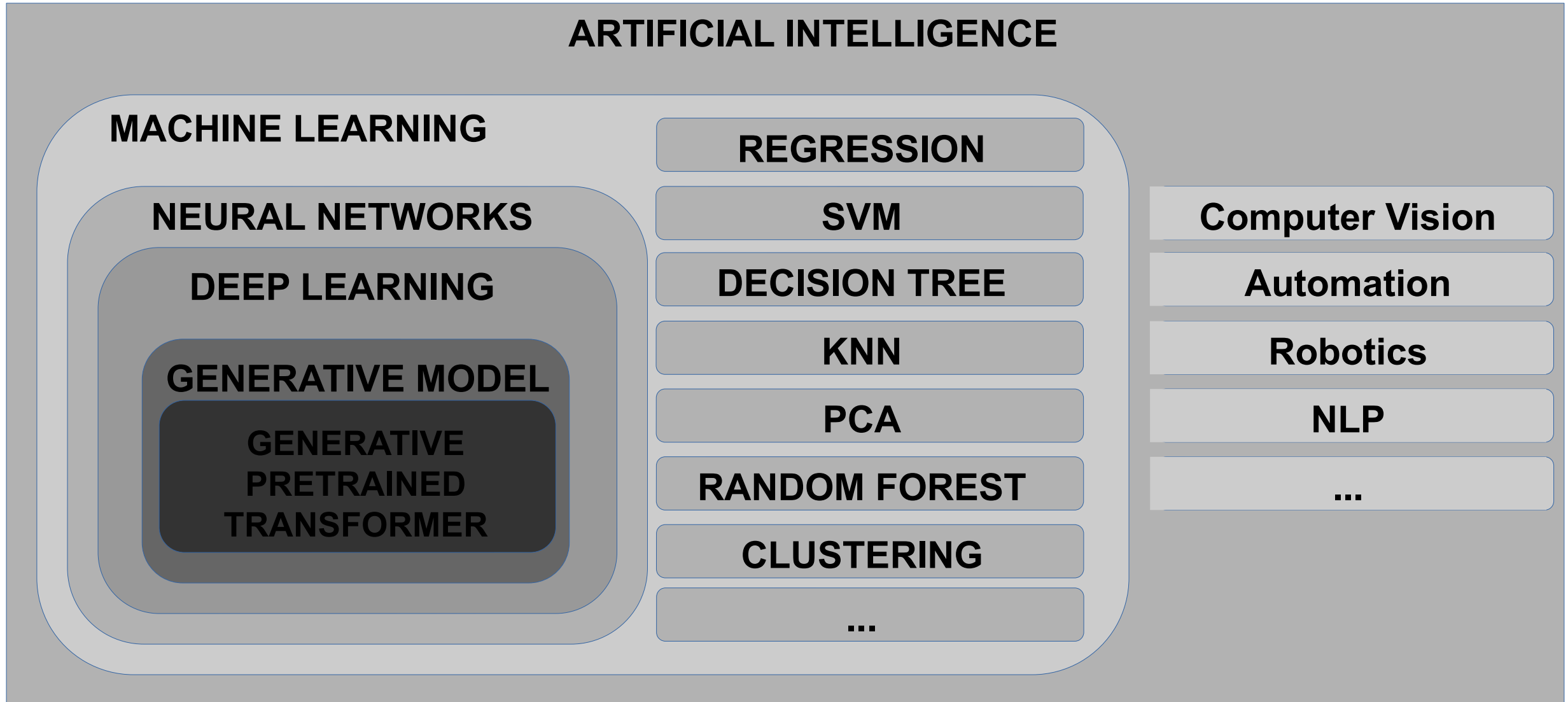
$O(n)$

- **Programming Language** : C/Rust programs run definitely faster than java/python programs **(for the same algorithm)**

# Quick Notes : ARTIFICIAL INTELLIGENCE

- Artificial intelligence (AI) is the capability of computational systems to perform tasks typically associated with human intelligence, such as learning, reasoning, problem-solving, perception, and decision-making. (Wikipedia)

# Quick Notes : ARTIFICIAL INTELLIGENCE



# Quick Notes : MACHINE LEARNING

**PROBLEM DEFINITION**

**DATA COLLECTION**

**DATA CLEANING**

**FEATURE EXTRACTION**

**TRAINING**

**VALIDATION**

**TESTING**

**DATA VISUALIZATION**

**REPORT GENERATION**

# Quick Notes : IMPLEMENTATION OF AI TASKS

**PYTHON**

**R**

**MATLAB**

**JAVA**

**C++**

**RUST**

**...**

**Homework and Project will be implemented in Python.**

**Python has large set of Machine Learning libraries that wraps native (C) libraries.**

**Python is easy to learn.**

- MEHMET PEKMEZCİ :
- ACADEMY :
- 2003 - BSc, Galatasaray University, Computer Engineering
- 2011 - MSc, Maltepe University, Computer Engineering
- 2025 - PhD, Gebze Technical University, Computer Engineering
- WORK :
- 2003 – 32BIT LTD. – Java Developer
- 2008 – HAVELSAN - Unix/Linux System Admin
- 2011 - HAVELSAN - Development Environment Admin
- 2017 – HAVELSAN - System Admin Team Leader
- 2019 – HAVELSAN - Infrastructure Group Leader (Sysops, Devops, Infra Teams)



# Introduce yourself

- Your name
- Your department
- Your expectations from the course
- Your skill of AI knowledge

# Course Overview

# Why Data Structures and Algorithms

- Data structures are vital in handling and manipulating large datasets in fields like machine learning and data science.
- Algorithms are used for data analysis, pattern recognition, and more.
- Code efficiency, e.g., optimizations in searching and sorting, which is critical for software performance.
- Structural approach to problem solving.
- Critical Thinking and Problem-Solving Skills: Encourages to break down complex problems into manageable components.
- Resource management, i.e., efficient use of system resources like memory and processing power.

# Purpose and Outcomes

Purpose of the course    This is a course designed to *enhance advanced Python programming, data structures, and algorithm skills* necessary for developing software, coding, and conducting group work in the field of data science and artificial intelligence

Learning outcomes    

- Ability to perform basic data structures and algorithm design and analysis
- Ability to collaborate in software development
- Possessing skills in current software development technologies

# Methodology

- Face-to-face lectures,
- Active participation expected,
- Hands-on coding,
- Homeworks: Individual work expected. **A pdf report and a python file is expected to be delivered.**
- Project: May be group project based on the project idea. **A pdf report is expected after each phase, and a python code tar ball is expected at the end of the semester.** Phases are indicated in “Project Phases” , report dates are indicated in “Schedule” pages.
- Two quizzes,
- One midterm exam,
- One final exam.

# Project Phases

- Data Collection
- Data Extraction (Standardized Format)
- Data Cleaning
- Find and replace missing value
- Normalize data
- Data Visualization
- Feature Extraction
- Training
- Validation
- Testing

# Project Report Templates

- <https://github.com/mehmetpekmezci/GTU-ENG-346/tree/main/Reports>

# Grading

|                     |            |
|---------------------|------------|
| <b>Homework</b>     | <b>20</b>  |
| <b>Quiz</b>         | <b>20</b>  |
| <b>Midterm Exam</b> | <b>20</b>  |
| <b>Project</b>      | <b>30</b>  |
| <b>Final Exam</b>   | <b>30</b>  |
| <b>Total</b>        | <b>130</b> |



# Schedule

| Week #  | Topic                           | Assessment   |
|---------|---------------------------------|--|
| Week 1  | Basic Python Concepts with GUI  |  |
| Week 2  | Matrix Operations and Numpy     | HW1  |
| Week 3  | Basic Bash Concepts             | <b>Project proposals</b>                                     |
| Week 4  | Data Analysis and Visualization | HW2 / <b>Project-Data-Collection-Report</b>                  |
| Week 5  | Object-Oriented Programming     | QUIZ1 / <b>Project-Data-Extraction-Report</b>                |
| Week 6  | Complexity                      | <b>Project-Data-Cleaning-And-Missing-Value-Report</b>        |
| Week 7  | Arrays, Linked Lists, Maps,     | HW3 / <b>Project-Data-Normalization-Visualization-Report</b> |
| Week 8  | Recursion                       | Midterm exam   |
| Week 9  | Stacks and Queues               | HW4 - <b>Project-Feature-Extraction-And-Training-Report</b>  |
| Week 10 | Linked Lists, Trees             | <b>Project-Validation-And-Testing-Report</b>                 |
| Week 11 | Graphs and NetworkX             | HW5  |
| Week 12 | Graphs and NetworkX             | QUIZ2  |
| Week 13 |                                 | HW6  |
| Week 14 | Final Projects                  |  |
| Week 15 | -                               |  |
| Week 16 | Final exam                      | Final project presentations                                  |

# Windows Subsystem for Linux (WSL)

```
## https://learn.microsoft.com/en-us/windows/wsl/install
```

```
## https://ubuntu.com/desktop/wsl
```

```
## In powershell :
```

```
Wsl --list --online
```

```
Wsl --install Ubuntu-24.04
```

```
Wsl -d Ubuntu-24.04
```

```
## Linux Shell Starts Here
```

```
sudo apt update
```

```
sudo apt install python3-pip
```

```
pip3 install matplotlib numpy pandas
```

```
## The first time you run code from Ubuntu, it will trigger a download of the necessary dependencies:
```

```
code .
```

# Textbooks and Other Resources

## Textbooks

- Goodrich, Michael T., Roberto Tamassia, and Michael H. Goldwasser, *Data structures and algorithms in Python*, John Wiley & Sons Ltd, 2013.
- Grus, Joel, *Data science from scratch: first principles with python*, O'Reilly Media, 2019.

## Recommended

- <https://www.kaggle.com/learn>
- <https://www.coursera.org/learn/python-data>
- <https://www.coursera.org/learn/python-data-analysis>

# Office Hours and Communication

- No office. We can talk between and after the lectures.
- Reach me at [mpekmezci@gtu.edu.tr](mailto:mpekmezci@gtu.edu.tr)