# ENCS5341 Machine Learning and Data Science

# Introduction

STUDENTS-HUB.com

Uploaded By: Jibreel Bornat

### What is Machine Learning?

 "Learning is any process by which a system improves performance from experience." - Herbert Simon

• *"Machine Learning is the field of study that gives computers the ability to learn without being explicitly programmed."* -*Arthur Samuel* 

 "A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E." - Tom Mitchell

A well-defined learning task is given by <P, T, E>.

STUDENTS-HUB.com

Uploaded By: Jibreel<sup>1</sup>Bornat

# Defining the Learning Task

#### A checkers learning problem:

- T: Playing checkers
- P: Percentage of games won against an arbitrary opponent
- E: Playing practice games against itself

#### A handwriting recognition learning problem:

- T: Recognizing hand-written words within images
- P: Percentage of words correctly classified
- E: Database of human-labeled images of handwritten words 10

#### A robot driving learning problem:

- T: Driving on four-lane highways using vision sensors
- P: Average distance traveled before a human-judged error
- E: A sequence of images and steering commands recorded while observing a human driver.







#### Uploaded By: Jibreel<sup>2</sup>Bornat

### Machine Learning vs. Traditional Programming

• Traditional Programming



• Machine Learning



STUDENTS-HUB.com

Uploaded By: Jibreel<sup>3</sup>Bornat

### When Do We Use Machine Learning?

ML is used when:

- Human expertise does not exist (Navigating on Mars) •
- Humans can't explain their expertise (Face Detection) •
- Models must be customized (Recommendation Systems) ullet
- Models are based on huge amounts of data (Genomics)  ${\color{black}\bullet}$









Data Science for Busi What You Need to Know about Data Mining and 



STUDENTS-HUB.com

Uploaded By: Jibreel<sup>4</sup>Bornat

## The Goal of Machine Learning

- Machine learning can be seen as learning to approximate functions from examples to make predictions about future unseen examples.
- Example: Learn a function to map images of animals to its category



Dogs STUDENTS-HUB.com



#### Unseen Example



Cat or Dog?

Uploaded By: Jibreel <sup>5</sup>Bornat

### The Goal of Machine Learning

- Machine learning can be seen as learning to approximate functions from examples to make predictions about future unseen examples.
- *Warning*: we might never know the true underling function.
- Example: Learning polynomials f: N -> N





# Types of Learning

- Supervised Learning
  - Given: training data + desired outputs (labels)

- Unsupervised Learning
  - Given: training data (without desired outputs)

- Reinforcement Learning
  - Rewards from sequence of actions



STUDENTS-HUB.com

#### Overview



STUDENTS-HUB.com

Uploaded By: Jibreel<sup>®</sup>Bornat

**Machine Translation** 



STUDENTS-HUB.com

Uploaded By: Jibreel<sup>9</sup>Bornat

Automatic Speech Recognition (ASR)



STUDENTS-HUB.com

Uploaded By: Jibree<sup>1</sup><sup>o</sup>Bornat

#### Weather Forecasting



STUDENTS-HUB.com

Uploaded By: Jibree<sup>1</sup>Bornat

#### **Recommendation Systems**



Customers who viewed this item also viewed these products



Dualit Food XL1500

Processor

\$560

📮 Add to cart



Kenwood kMix Manual

Espresso Machine

\*\*\*\*\*

\$250

Select options

🕁 Add to cart



Weber One Touch Gold NoMU Premium Charcoal Spi Grill-57cm

\$225

NoMU Salt Pepper and Spice Grinders \$3

😾 View options

#### STUDENTS-HUB.com

Uploaded By: Jibree<sup>12</sup>Bornat

Optical Character Recognition (OCR)



[LeCun et al. 1998]

STUDENTS-HUB.com

Uploaded By: Jibree<sup>13</sup>Bornat

#### **Fraud Detection**



STUDENTS-HUB.com

Uploaded By: Jibree<sup>14</sup>Bornat

#### Image Classification



STUDENTS-HUB.com

howler monkey ploaded By: Jibreel Bornat

#### Scene Understanding



STUDENTS-HUB.com

Uploaded By: Jibree<sup>16</sup>Bornat

#### Data Science

- Data science is a field of study and practice that involves the collection, storage, and processing of data in order to derive important insights into a problem or a phenomenon.
- Data may be generated by humans (surveys, logs, etc.) or machines (weather data, road vision, etc.), and could be in different formats (text, audio, video, augmented or virtual reality, etc.)

Uploaded By: Jibree<sup>†7</sup>Bornat

#### Data Science



Image from javatpoint

STUDENTS-HUB.com

Uploaded By: Jibree<sup>18</sup>Bornat



STUDENTS-HUB.com

Uploaded By: Jibree<sup>19</sup>Bornat

- **Phase 1—Discovery**: In Phase 1, the team learns the business domain, including attempted similar projects in the past from which they can learn. The team assesses the resources available to support the project in terms of people, technology, time, and data. Important activities in this phase include framing the business problem as an analytics challenge that can be addressed in subsequent phases and formulating initial hypotheses to test and begin learning the data.
- Phase 2—Data preparation: Phase 2 the team works with data and perform analytics for the duration of the project. The team needs to execute extract, load, and transform (ELT) or extract, transform and load (ETL) to get data into the sandbox. The ELT and ETL are sometimes abbreviated as ETLT. Data should be transformed in the ETLT process so the team can work with it and analyze it. In this phase, the team also needs to familiarize itself with the data thoroughly and take steps to condition the data.

Uploaded By: Jibree<sup>1</sup><sup>®</sup>Bornat

- **Phase 3—Model planning**: Phase 3 is model planning, where the team determines the methods, techniques, and workflow it intends to follow for the subsequent model building phase. The team explores the data to learn about the relationships between variables and subsequently selects key variables and the most suitable models.
- **Phase 4—Model building**: In Phase 4, the team develops datasets for testing, training, and production purposes. In addition, in this phase the team builds and executes models based on the work done in the model planning phase. The team also considers whether its existing tools will suffice for running the models, or if it will need a more robust environment for executing models and workflows (for example, fast hardware and parallel processing, if applicable).

Uploaded By: Jibree<sup>1</sup>Bornat

- Phase 5—Communicate results: In Phase 5, the team, in collaboration with major stakeholders, determines if the results of the project are a success or a failure based on the criteria developed in Phase 1. The team should identify key findings, quantify the business value, and develop a narrative to summarize and convey findings to stakeholders.
- **Phase 6—Operationalize**: In Phase 6, the team delivers final reports, briefings, code, and technical documents. In addition, the team may run a pilot project to implement the models in a production environment.

Uploaded By: Jibree<sup>2</sup>Bornat

### Data Scientist vs. Machine Learning Engineer

- Data scientist and ML engineer are two closely related roles in the field of data science and machine learning.
- Here is a brief comparison between their roles
  - Data Scientists:
    - Focus on data analysis, i.e., collecting, cleaning, and analyzing data to extract insights and generate actionable recommendations. Apply exploratory data analysis (EDA) to gain a deep understanding of the data and its implications. Use statistical methods and data visualization to explore and communicate their findings.
  - ML Engineers:
    - ML engineers focus on designing, building, and deploying machine learning models into production systems. They are responsible for optimizing machine learning models for performance, efficiency, and scalability. They often work on model fine-tuning and hyperparameter optimization.

STUDENTS-HUB.com

Uploaded By: Jibree<sup>73</sup>Bornat



#### Regression



STUDENTS-HUB.com

Uploaded By: Jibreef<sup>5</sup>Bornat

#### Classification



STUDENTS-HUB.com

Uploaded By: Jibreef<sup>6</sup>Bornat

Kernel methods and SVM





Right: Features (x,y,  $x^2+y^2$ )



Uploaded By: Jibreef<sup>7</sup>Bornat

#### Model selection / evaluation, Regularization



STUDENTS-HUB.com

Uploaded By: Jibreef<sup>8</sup>Bornat

**Ensemble methods** 





Uploaded By: Jibreef<sup>9</sup>Bornat

#### STUDENTS-HUB.com



HMM

Modelling sequential data



STUDENTS-HUB.com

Uploaded By: Jibreel<sup>®</sup>Bornat

#### Neural Networks, Deep Learning









Uploaded By: Jibreel<sup>31</sup>Bornat

STUDENTS-HUB.com



STUDENTS-HUB.com

Uploaded By: Jibree<sup>2</sup>Bornat

**Dimensionality Reduction and Autoencoders** 





Uploaded By: Jibreel<sup>3</sup>Bornat

STUDENTS-HUB.com