CONESTOGA

Connect Life and Learning

CSCN 8000

Artificial Intelligence Algorithms and Mathematics

Outline

- ✓ General Introductions
- ✓ Review to Course Policies and Procedures-Applicable
- ✓ Introduction to AI
- ✓ About ML
- ✓ Scope of ML
- ✓ Machine Learning Project Approach-checklist
- ✓ Essential Technical Tools
- ✓ Case studies and Applications for AI

Learning Objectives

- ✓ Understand Course Policies and Procedures-Applicable.
- ✓ Comprehend the difference between AI and ML.
- ✓ List the different types of machine learning techniques.
- ✓ Understand the different feature types in ML systems.
- ✓ Discuss the strengths and weaknesses of ML.
- ✓ Understand the technical tools available in the course for implementation of ML algorithms.

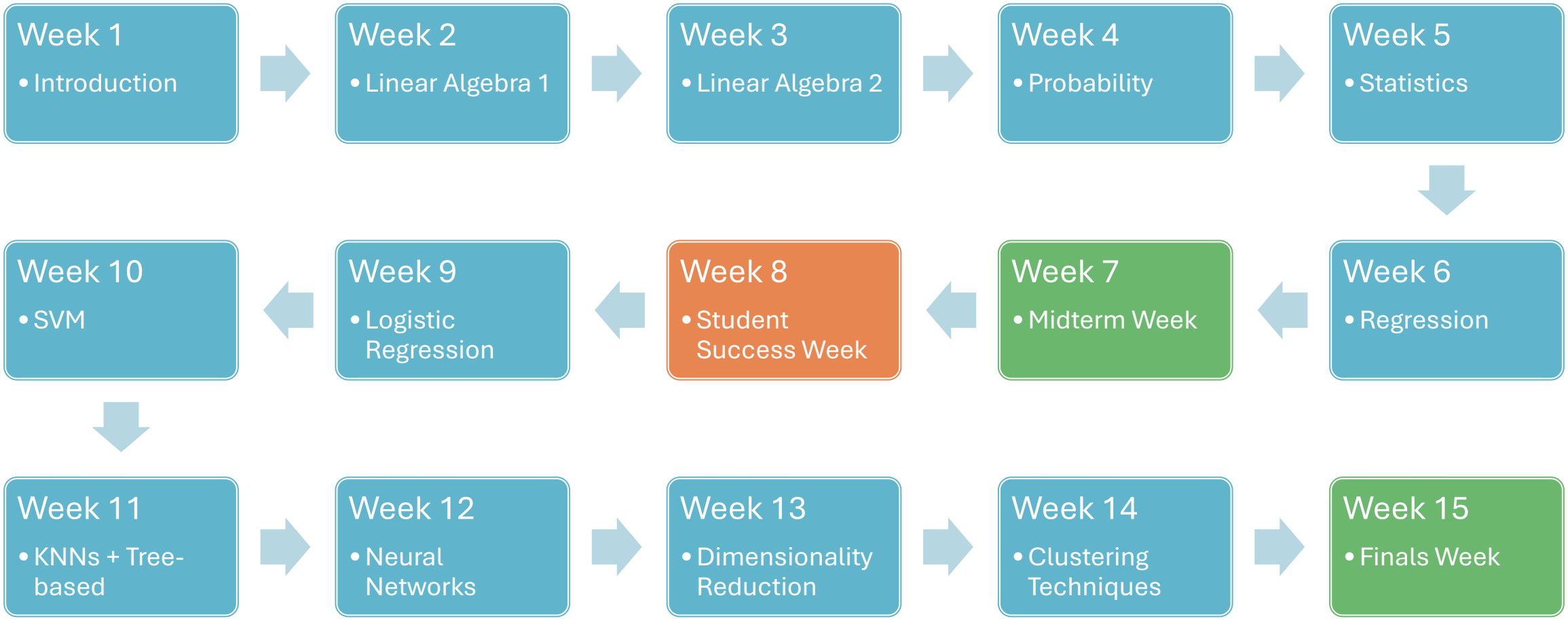
Program Handbook

- https://www.conestogac.on.ca/handbook/1557

Resources

- **Practical Statistics for Data Scientists, 2nd Edition Peter Bruce, Andrew Bruce, Peter Gedeck**
- **Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow, 2nd Edition By Aurélien Géron**

Course Content



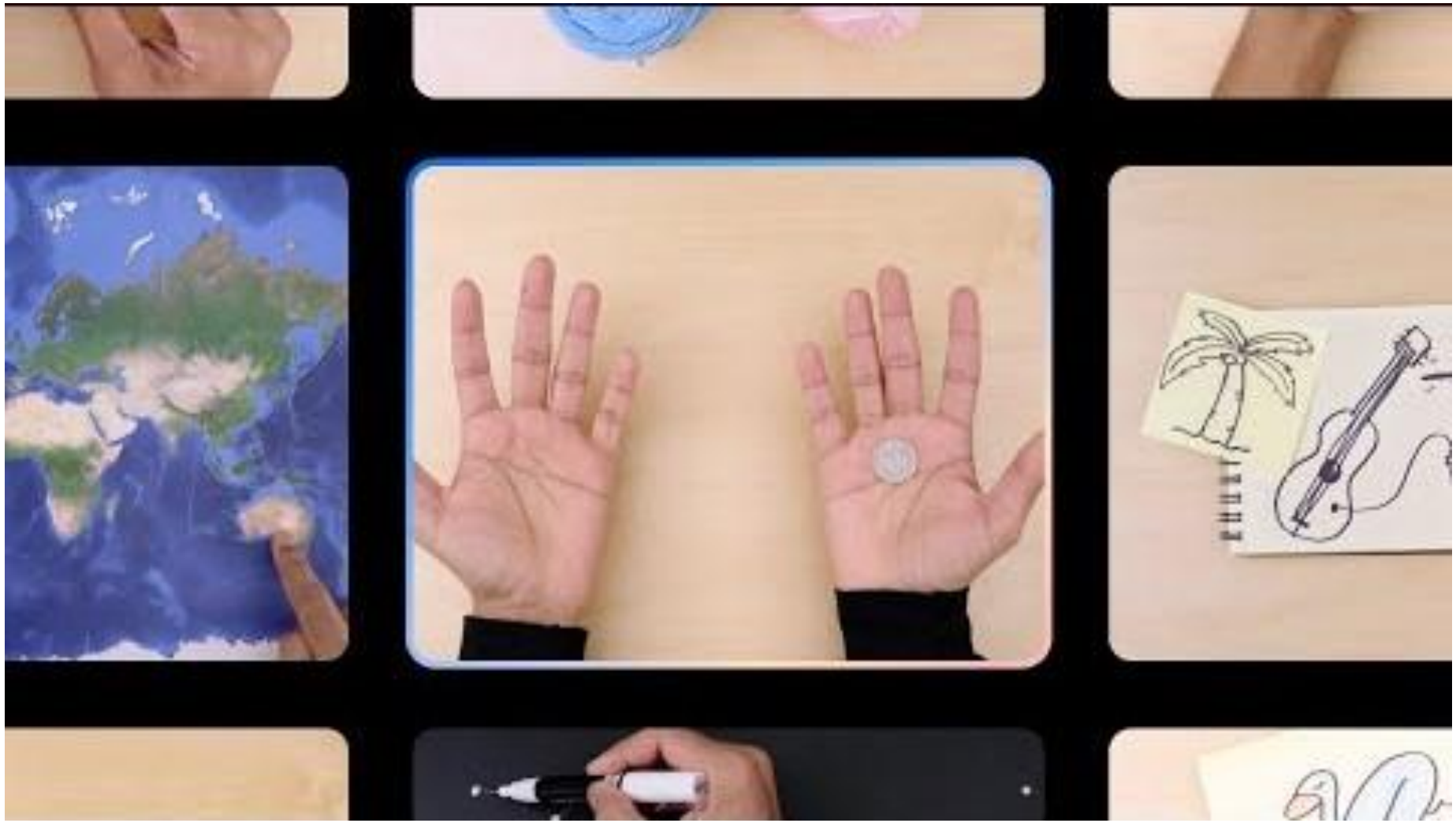
Course Deliverables

- 3 Labs → 10% each
- 3 Assignments → 10% each
- Midterm (20%):
  - Midterm Part A: Theoretical → 10%
  - Midterm Part B: Practical → 10%
- Final (20%):
  - Final Part A: Theoretical → 10%
  - Final Part B: Practical → 10%

Online Session Rules

- Attendance is mandatory and highly encouraged to properly understand the complex mathematical topics and solve the assignments/exams.
- Cameras should always be turned on unless prior approval has been given for special circumstances.
- The online sessions are supposed to be interactive and collaborative so feel free to ask questions or open discussions whenever possible.
- All sessions will be recorded. By participating, you consent to being recorded.


Capture



Introduction to Artificial Intelligence

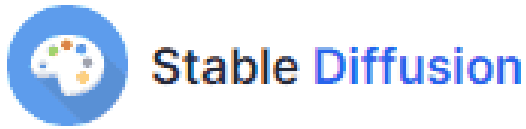
- Artificial Intelligence is an area of computer science, where the goal is to enable computers and machines to perform human like tasks and simulate human behavior.
- Machine Learning is subset of AI that tries to solve a specific problem and make predictions using data.
- Data Science is a field that attempts to find patterns and draw insights from the data.
- Mostly commonly driven tool in AI is Machine Learning

Categorization

- AI can be categorized into ANI and AGI
- ANI –Artificial Narrow Intelligence –Example : Self Driving Car, Virtual Assistants
- AGI- Mimicking the Human Ability

What is Machine Learning

- Subdomain of computer science that focusses on algorithms which help a computer learn **from data** without explicit programming



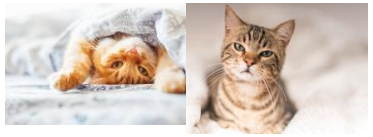
Types of Machine Learning

- Supervised Learning – Using Labelled Inputs – Outputs
  - Example : Identifying Pictures


CAT

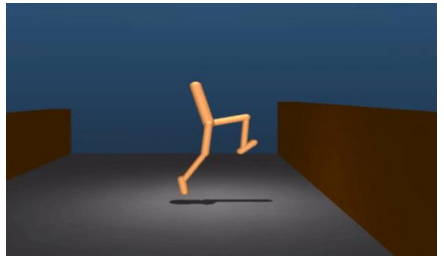
DOG

- Unsupervised Learning-Uses Unlabelled data to identify patterns in the data
  - Example : Clustering the Pictures


Structure 1

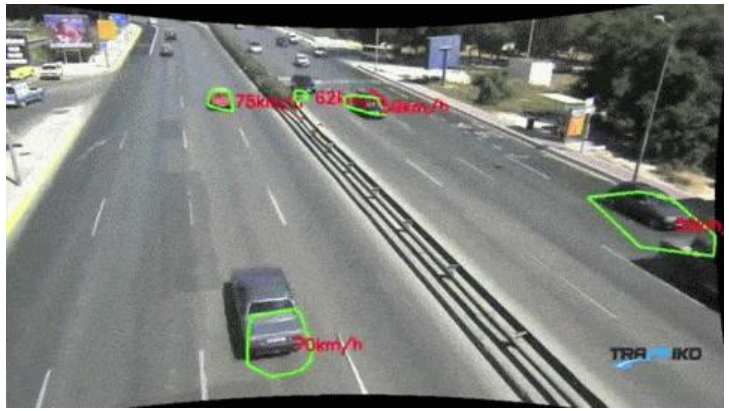

Structure 2

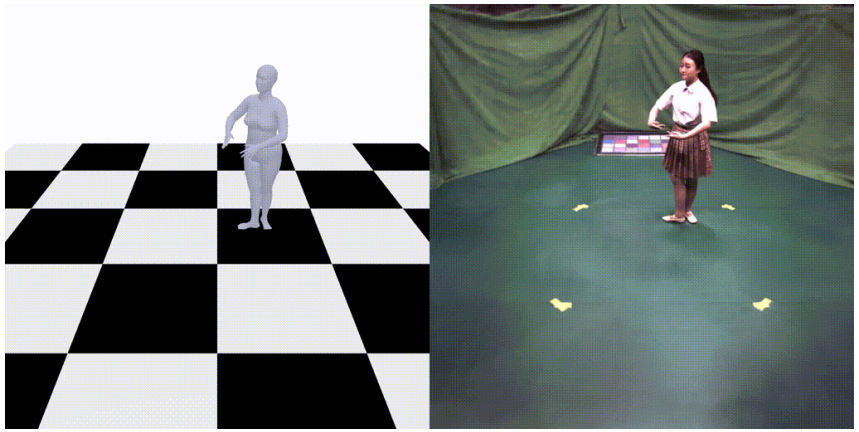
- Reinforcement Learning
  - Agent learning in interactive environment depending on rewards and penalties.



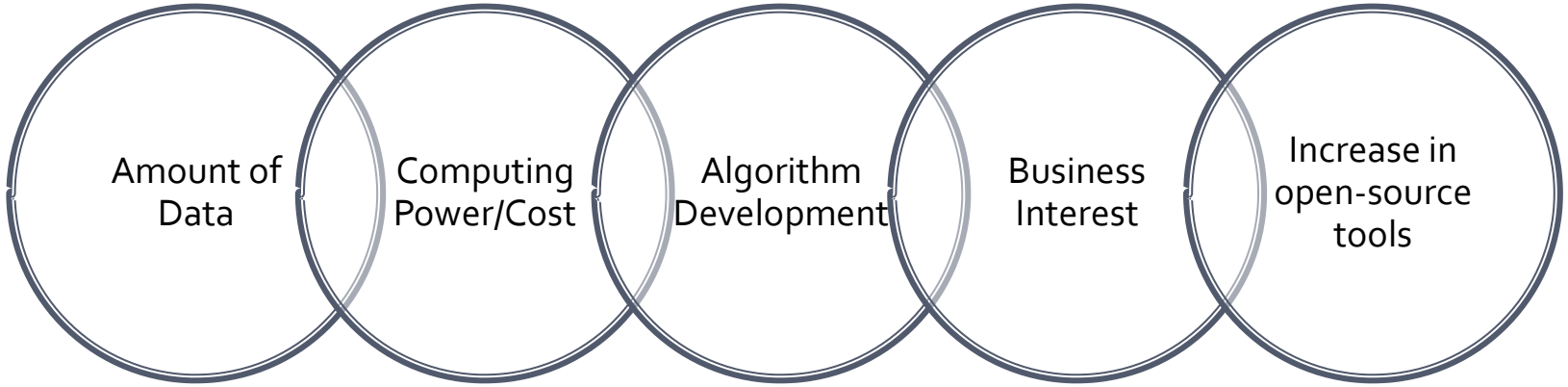
Supervised Learning Applications

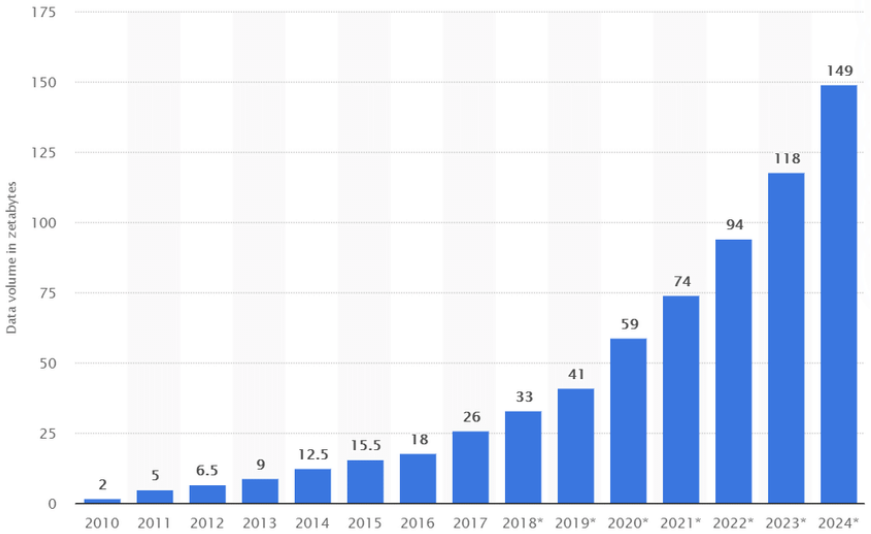
- Spam Filtering - >Input – email ,Output – Spam
- Speech Recognition -> Input – Audio ,Output – -Text Transcript
- Language Translation -> Input – English ,Output – Chinese
- Self-driving Car -> Input – image, radar ,Output – Position of the cars

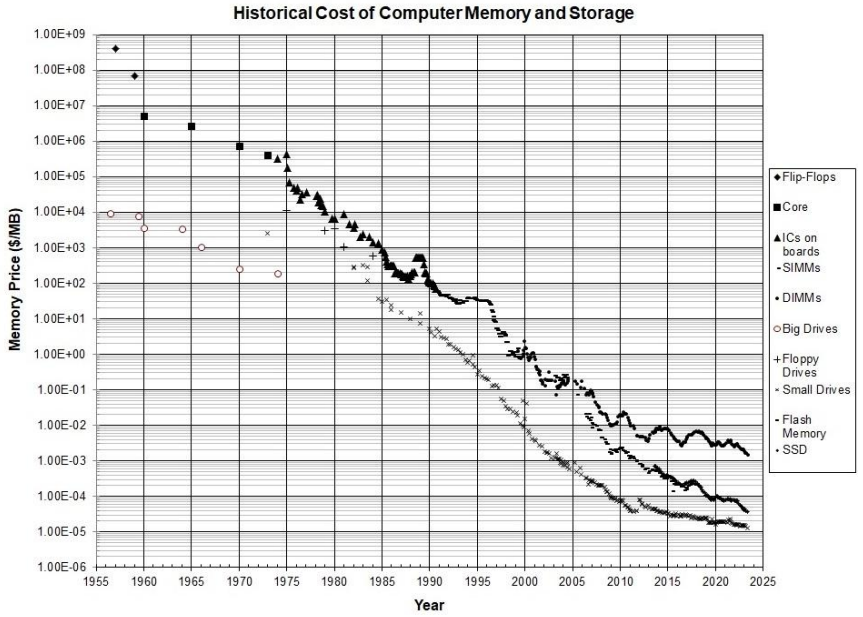




Why is ML booming?

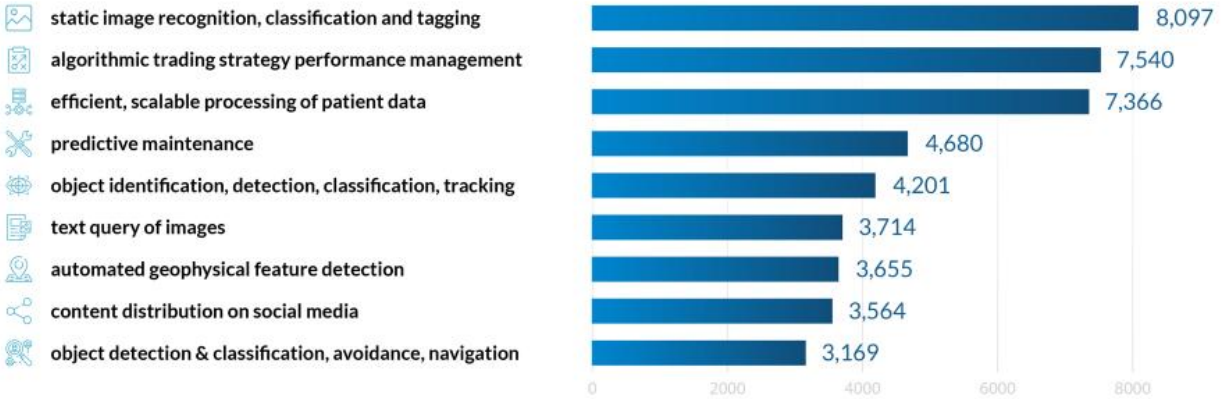






1 Global AI revenue forecast by 2025, ranked by use case in millions US dollar

Source: Statista

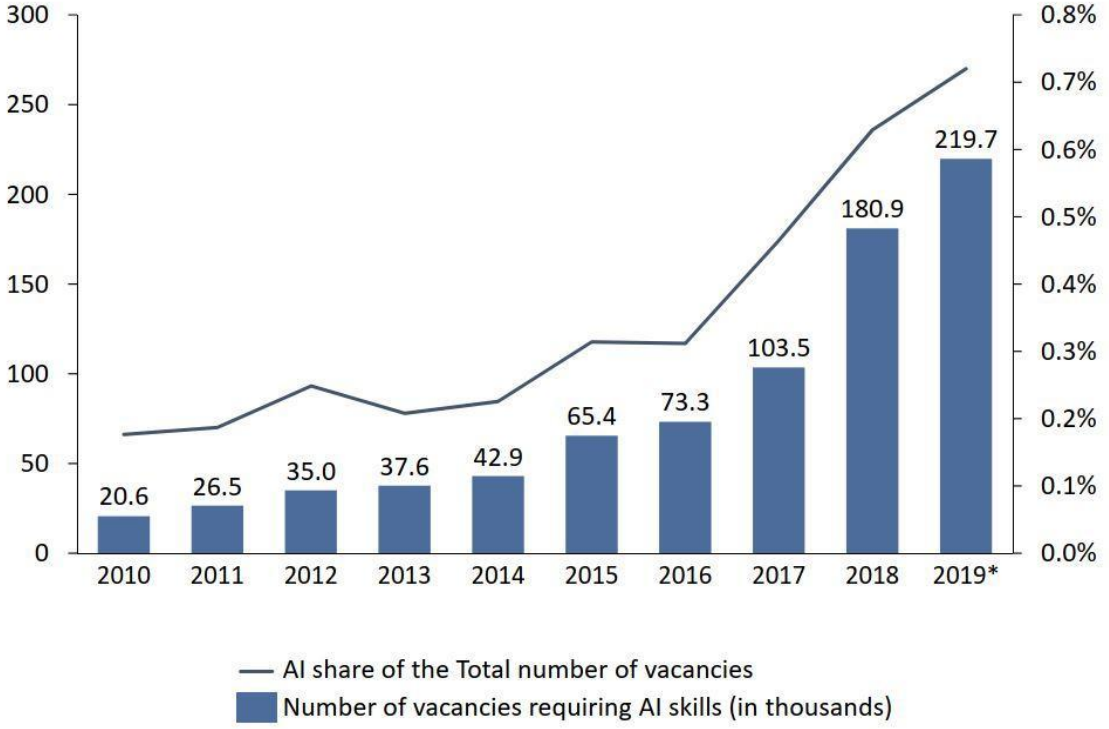


2 Penetration of artificial intelligence skills, by country

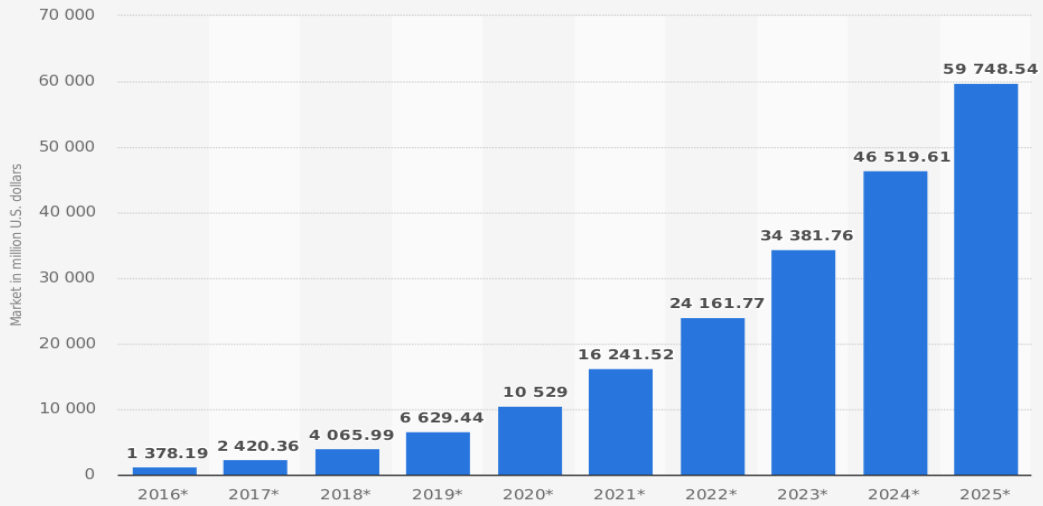
Source: Dun & Bradstreet

3 Organizations deploying AI, by functional areas

Source: Medium



Revenues from the artificial intelligence (AI) market worldwide, from 2016 to 2025 (in million U.S. dollars)

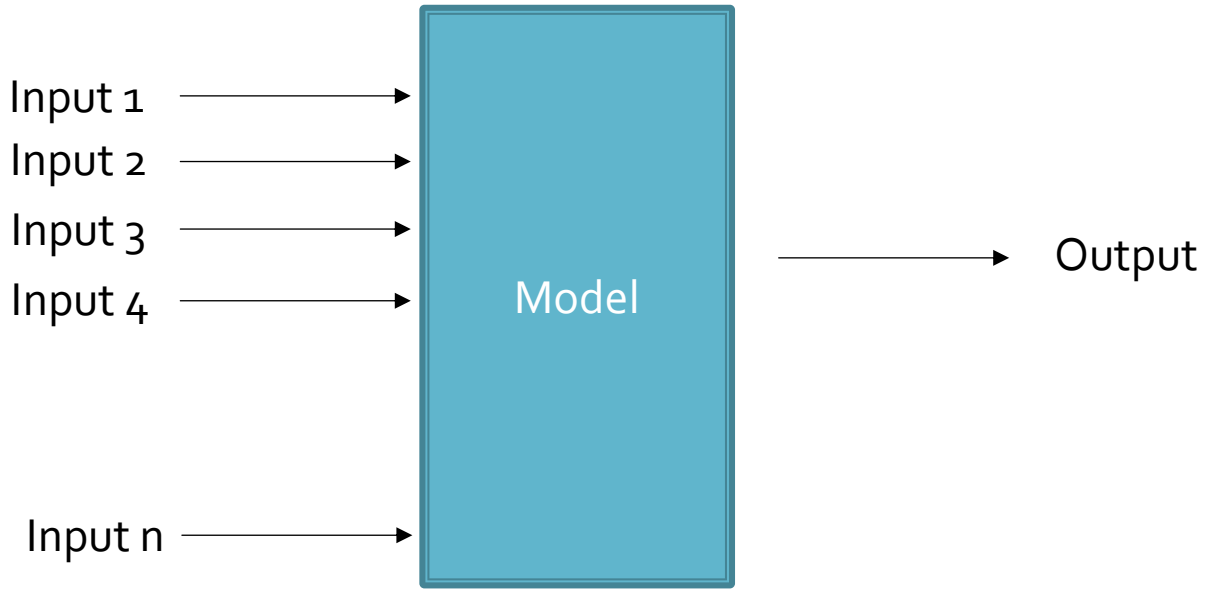


Source
Tractica
© Statista 2017

Additional Information:
Worldwide; 2017

statista

Machine Learning Model



Features: Qualitative Data Types

- Qualitative: Categorical data with finite number of categories.
  - Nominal Data: No inherent order
    - E.g. Red/Blue/Green
  - Ordinal Data: Posses meaningful order.
    - E.g. Excellent/Very Good/Good
  - Binary Data: Only two categorical values.
    - E.g. Yes/No
- One Hot Encoding: Used for processing by ML models.

UCI Machine Learning Repository: MAGIC Gamma Telescope Data Set

{USA,India ,Canada}

USA – { 1,0,0}
India–{ 0,1,0}
Canada-{ 0,0,1}

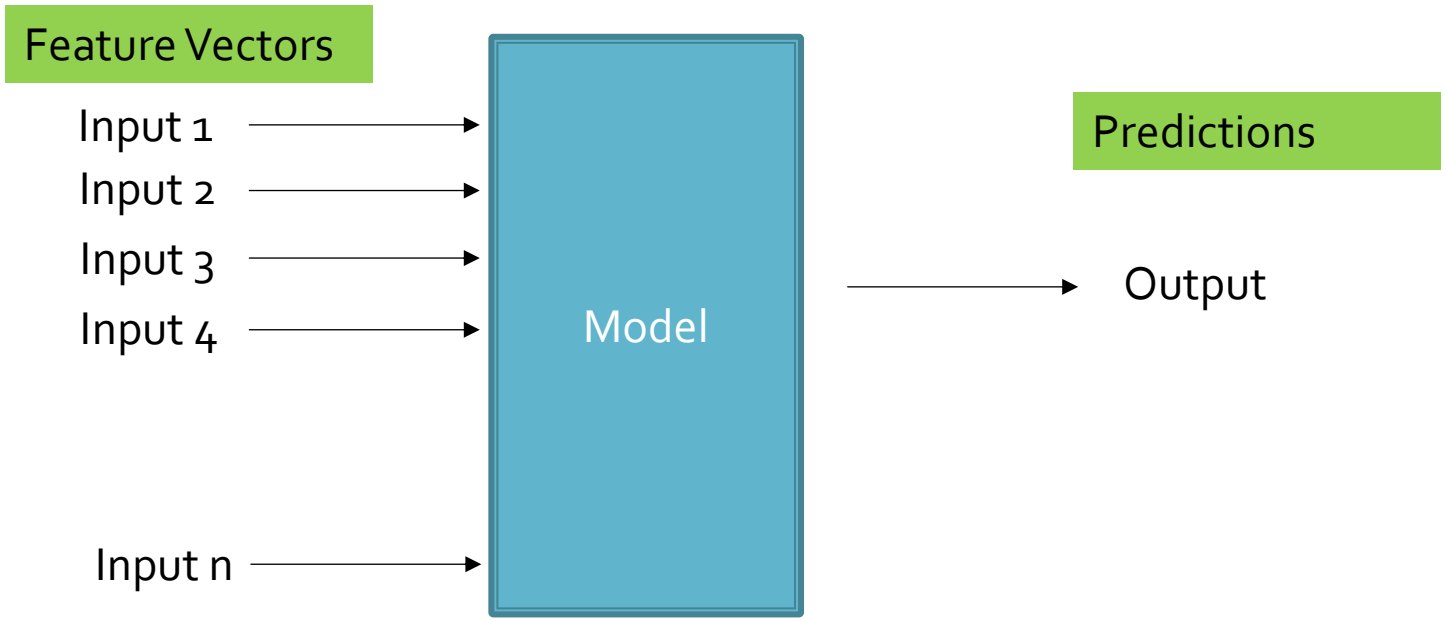
Features: Quantitative Data Types

- Quantitative –Numerical Values Data that could be either discrete or continuous
  - Discrete: distinct, separate values with clear gaps
    - Example: # of Employees
  - Continuous: can take on any value within a range
    - Ratio data: has a natural order **and** a meaningful zero point
      - Example: Age
    - Interval data: has a natural order **but lacks** a meaningful zero point
      - Example: Temperature measured in Celsius

Features: Q/A

- Bank account balance is an example of which type of data?
  - (a) Ratio
  - (b) Interval
- Which of the following is an example of nominal data?
  - a) Educational level
  - b) Types of fruits
- Is the measurement of pH (acidic or alkaline levels) considered which type of data?
  - a) Discrete
  - b) Ratio
  - c) Interval
- Which of the following is an example of interval data?
  - a) Height in centimeters
  - b) Weight in kilograms
  - c) Temperature in degrees Celsius
  - d) Distance in meters
- On a Likert scale (e.g., strongly disagree, disagree, neutral, agree, strongly agree), what type of data is obtained?
  - a) Ratio
  - b) Interval
  - (c) Ordinal

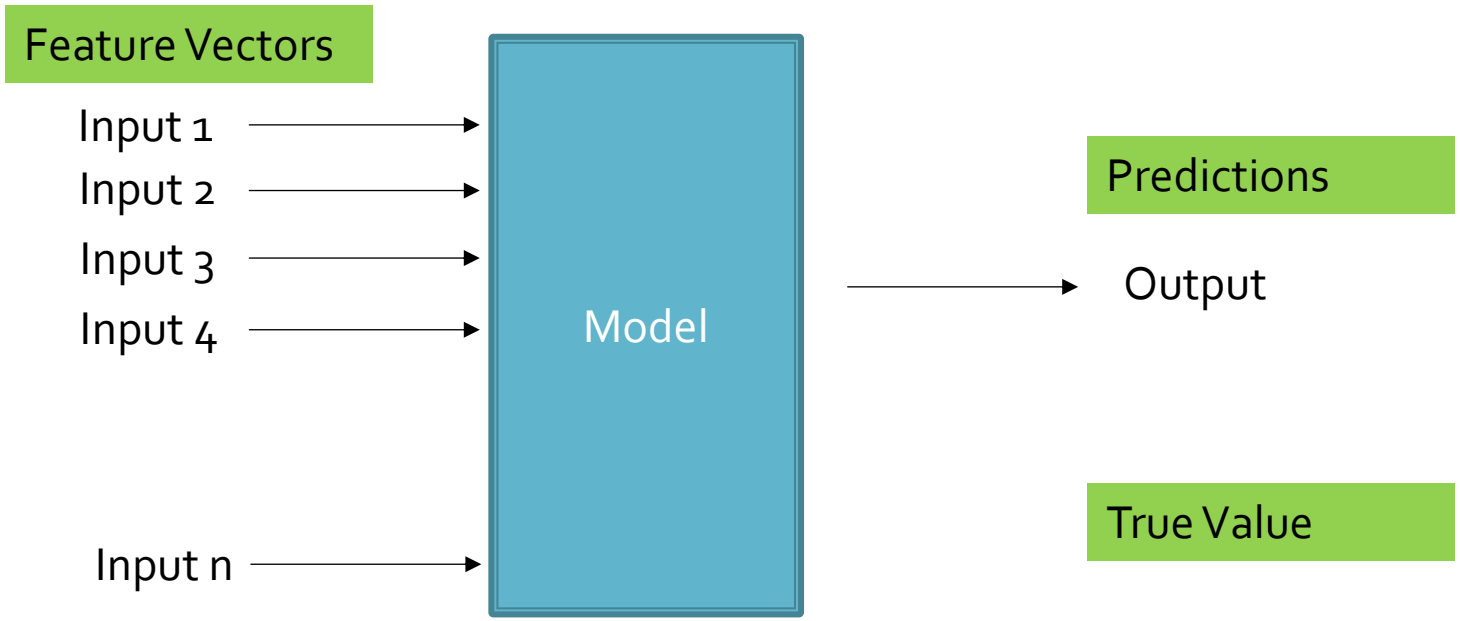
Machine Learning Model



Supervised Learning

- Classification : Predicting discrete classes /categories
  - Binary Classification(between two) ,Example : True/False, Object Identification- Cat /Not Cat
  - Multiclass classification(more than two) ,Example : Animal Species
  - Multilabel classification(more than two) ,Example : Colors in Image
- Regression :Predicting continuous values /Predicting the value that is close to the true value.
  - Examples: Price Prediction ,Temperature Prediction

Machine Learning Model



Data Set

- https://www.kaggle.com/c/titanic
- https://www.kaggle.com/datasets/sobhanmoosavi/us-accidents
- UCI Diabetes Data Set | Kaggle

ML Scope

Weaknesses	Commonsense Reasoning
	Contextual Understanding
	Causality and Explainability
	Labeled Data Dependence
	Ethical and Bias Issues
	Robustness and Security
	Privacy Concerns
	Continual Learning

Data Repositories

- Popular open data repositories:
  - OpenML.org
  - Kaggle.com
  - PapersWithCode.com
  - UC Irvine Machine Learning Repository
  - Amazon's AWS datasets
  - TensorFlow datasets
- Meta portals (they list open data repositories):
  - DataPortals.org
  - OpenDataMonitor.eu
- Other pages listing many popular open data repositories:
  - Wikipedia's list of machine learning datasets
  - Quora.com
  - The datasets subreddit

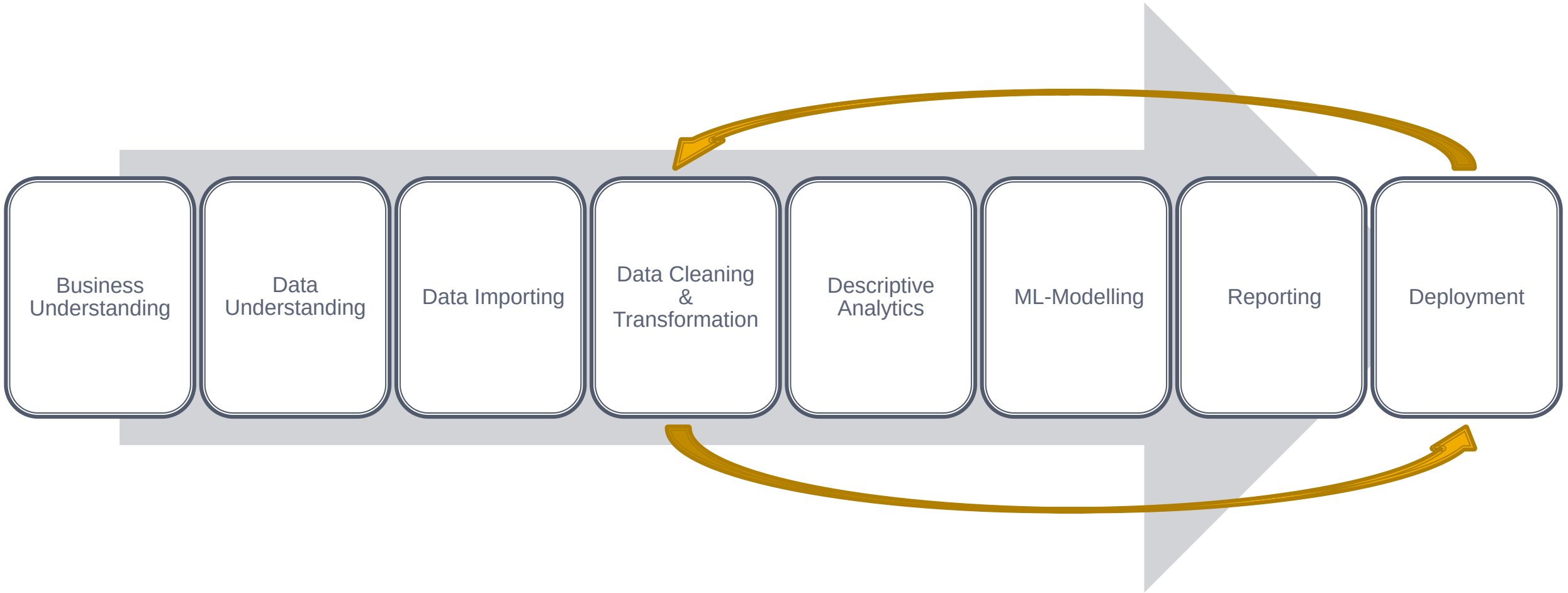
AI Technical Tools

- Machine Learning Frameworks
  - PyTorch
  - Tensorflow
  - Keras
  - MXNet
  - CNTK
  - Caffe
  - PaddlePaddle
  - Scikit-Learn
  - R
  - Weka
- Research Activity
  - Arxiv
- GitHub Repositories
  - Open Source

CPU Vs GPU

- CPU –Computer Processing Unit
- GPU :Graphics Processing Unit
- Cloud Vs On Premises

Machine Learning Project Checklist




Golden Stripes.


1984
air head
a short film
by shy kids

Thank you!

- Open Discussion

References

- Essential Scrum: A Practical Guide to the Most Popular Agile Process, Kenneth S. Rubin.
- Product Vision Board by Roman Pichler

Thank You

Youssef Abdelkareem

yabdelkareem@conestogac.on.ca

**** Credits for Dr. Anasuya Bhima for assisting with the slides content***