Lisbon School of Economics & Management
Universidade de Lisboa

LISBOA
UNIVERSIDADE DE LISBOA

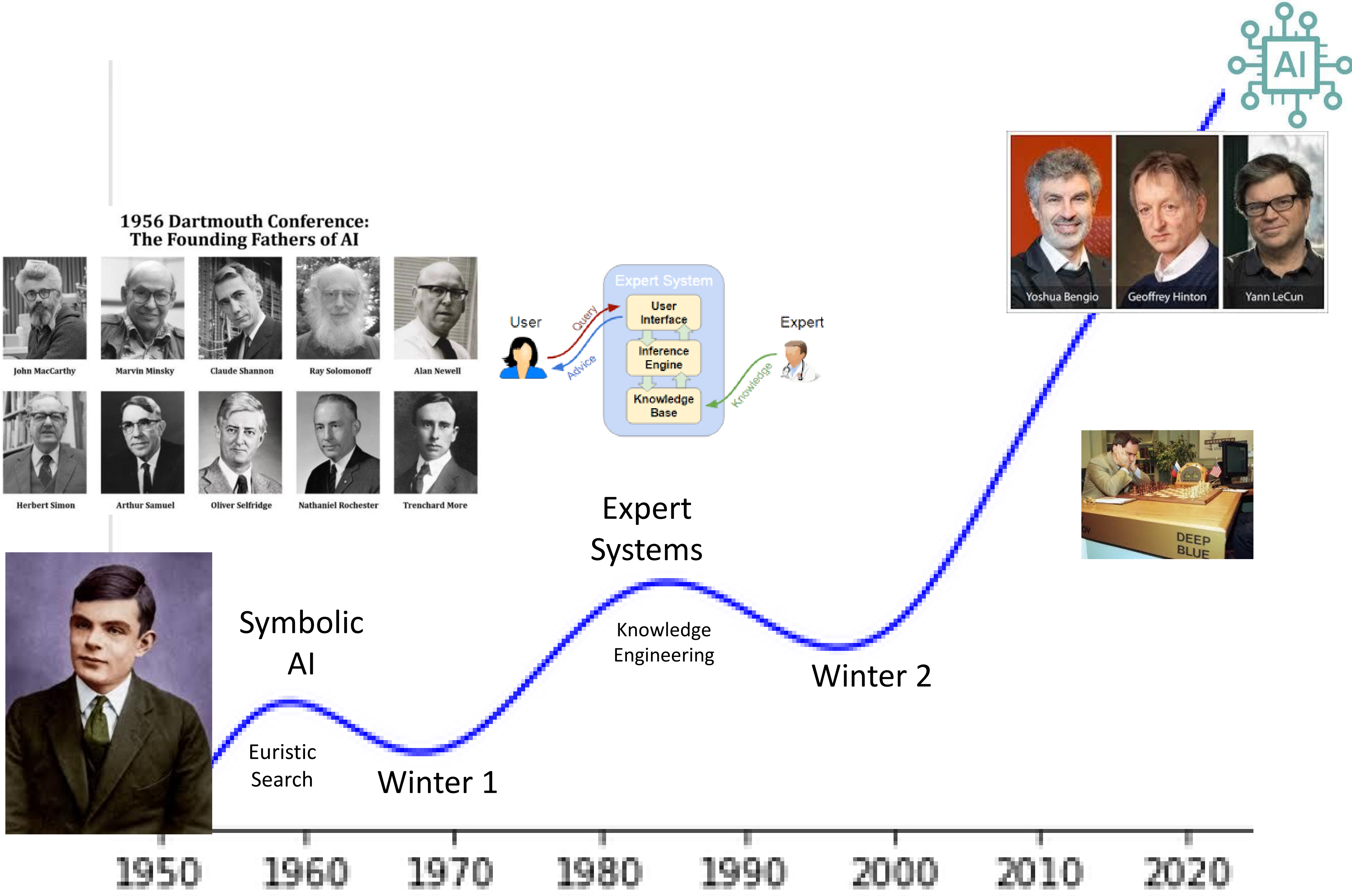
Carlos J. Costa

MACHINE LEARNING

WHAT IS A.I.?

Artificial Intelligence(AI)

- Artificial intelligence refers to the development of computer-based solutions that can perform tasks which mimic human intelligence.



1956 Dartmouth Conference:
The Founding Fathers of AI

Symbolic AI

Euristic Search

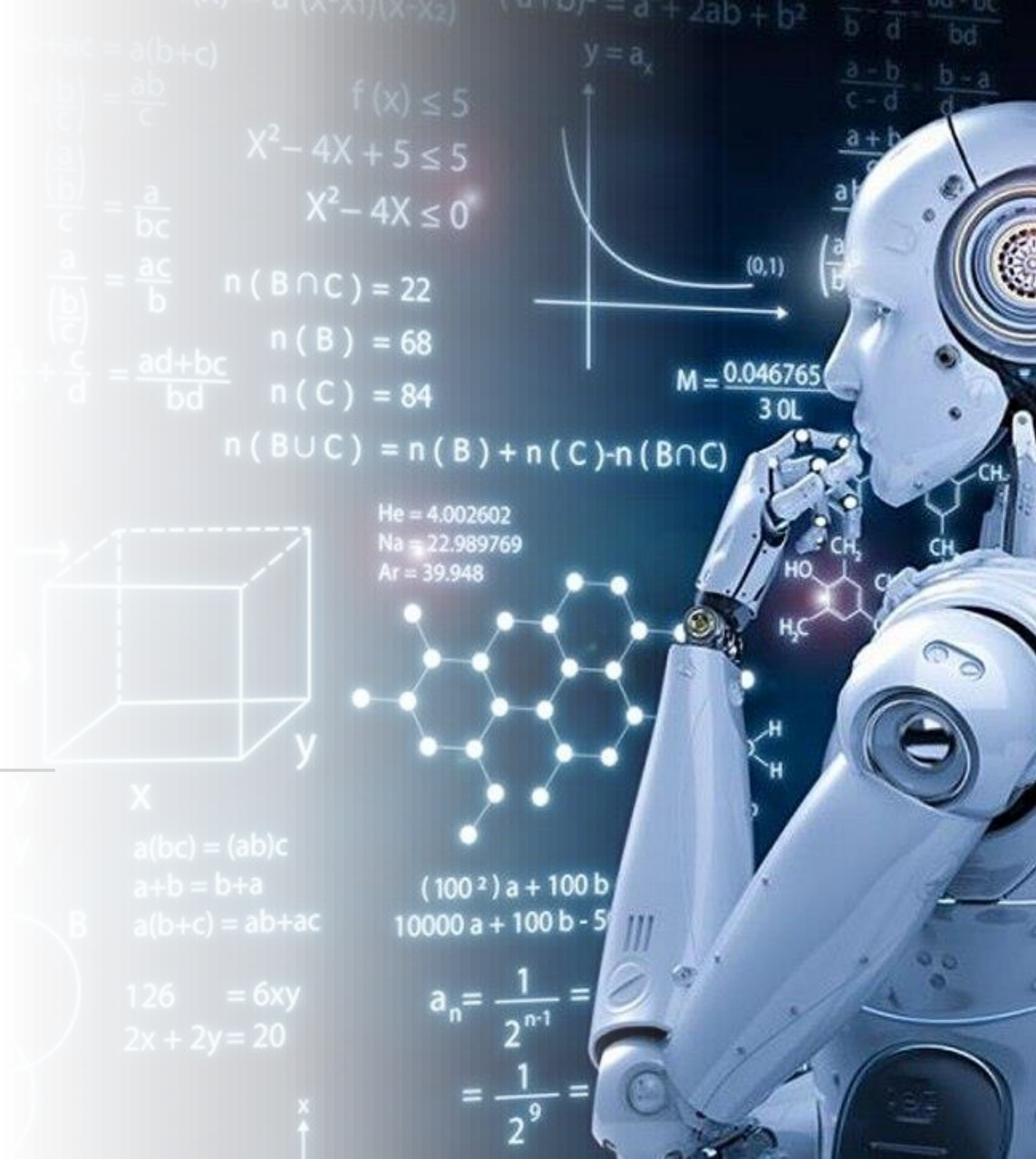
Winter 1

Expert Systems

Knowledge Engineering

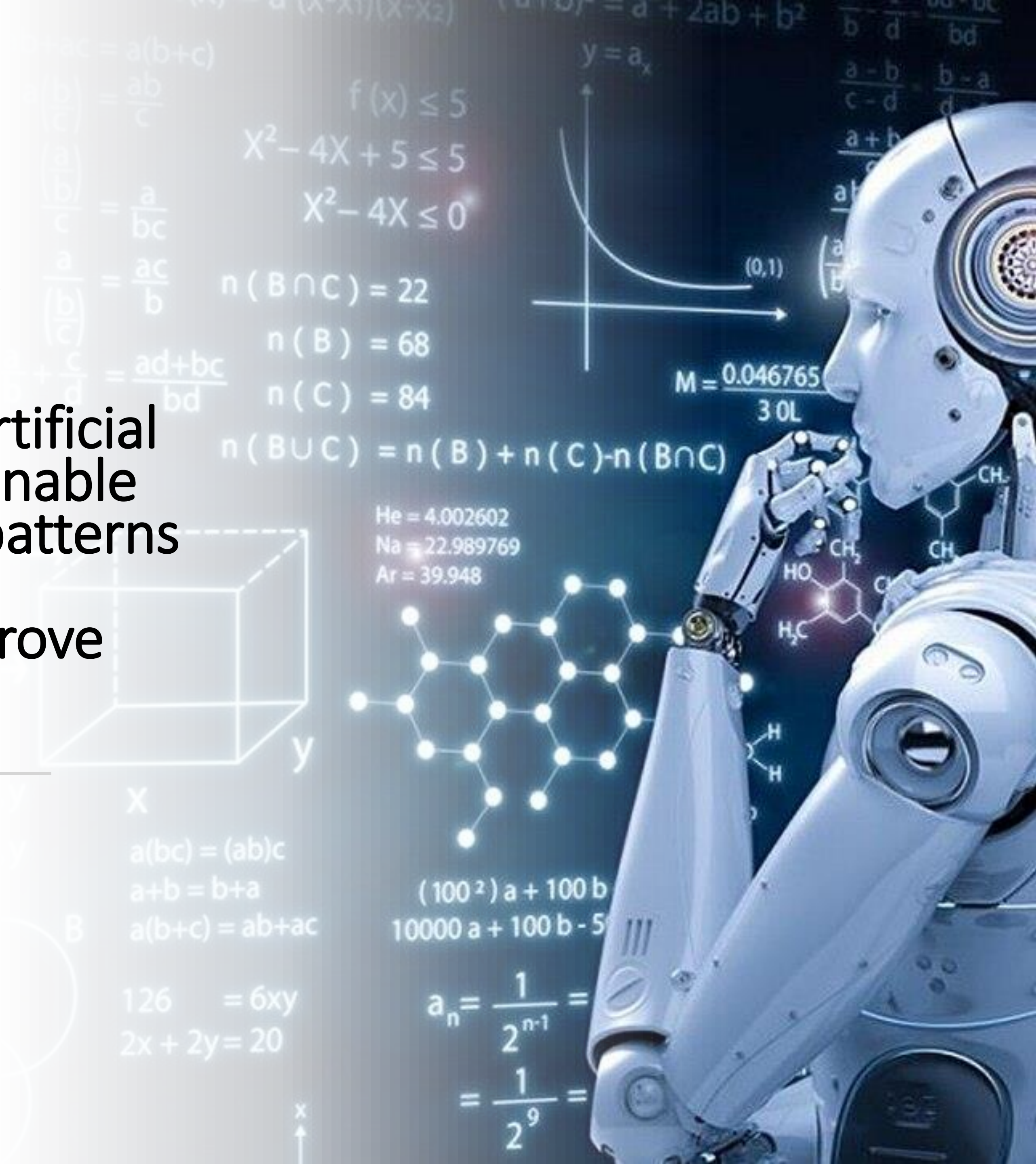
Winter 2

What is Machine Learning?

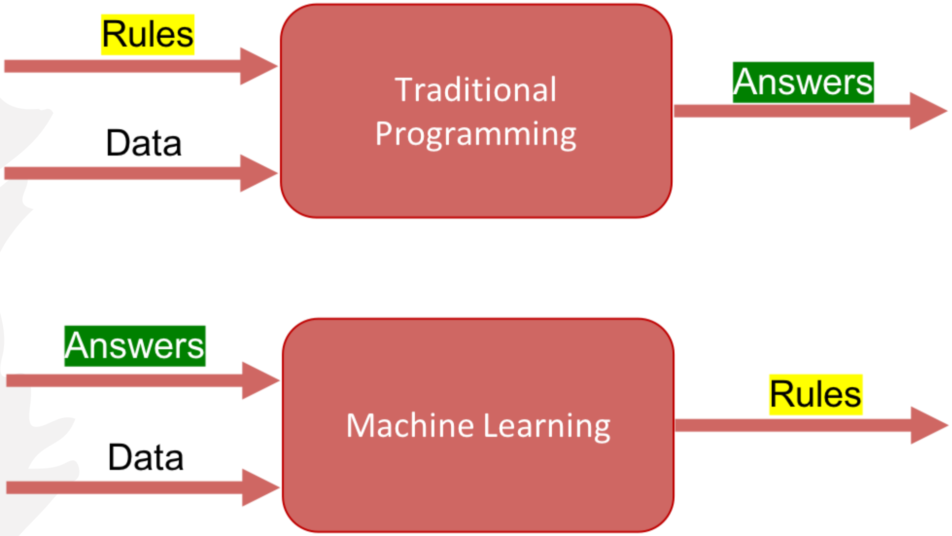


Machine Learning

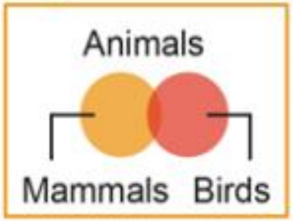
is as a subset of artificial intelligence that enable systems to learn patterns from data and subsequently improve from experience



Tradicional programming vs. Machine Learning



Symbolists



Use symbols, rules, and logic to represent knowledge and draw logical inference

Favored algorithm

Rules and decision trees

Bayesians

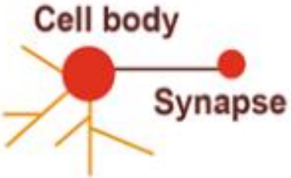
Likelihood
Prior
Posterior
Margin

Assess the likelihood of occurrence for probabilistic inference

Favored algorithm

Naive Bayes or Markov

Connectionists



Recognize and generalize patterns dynamically with matrices of probabilistic, weighted neurons

Favored algorithm

Neural networks

Evolutionaries

Generate variations and then assess the fitness of each for a given purpose

Favored algorithm

Genetic programs

Analogizers

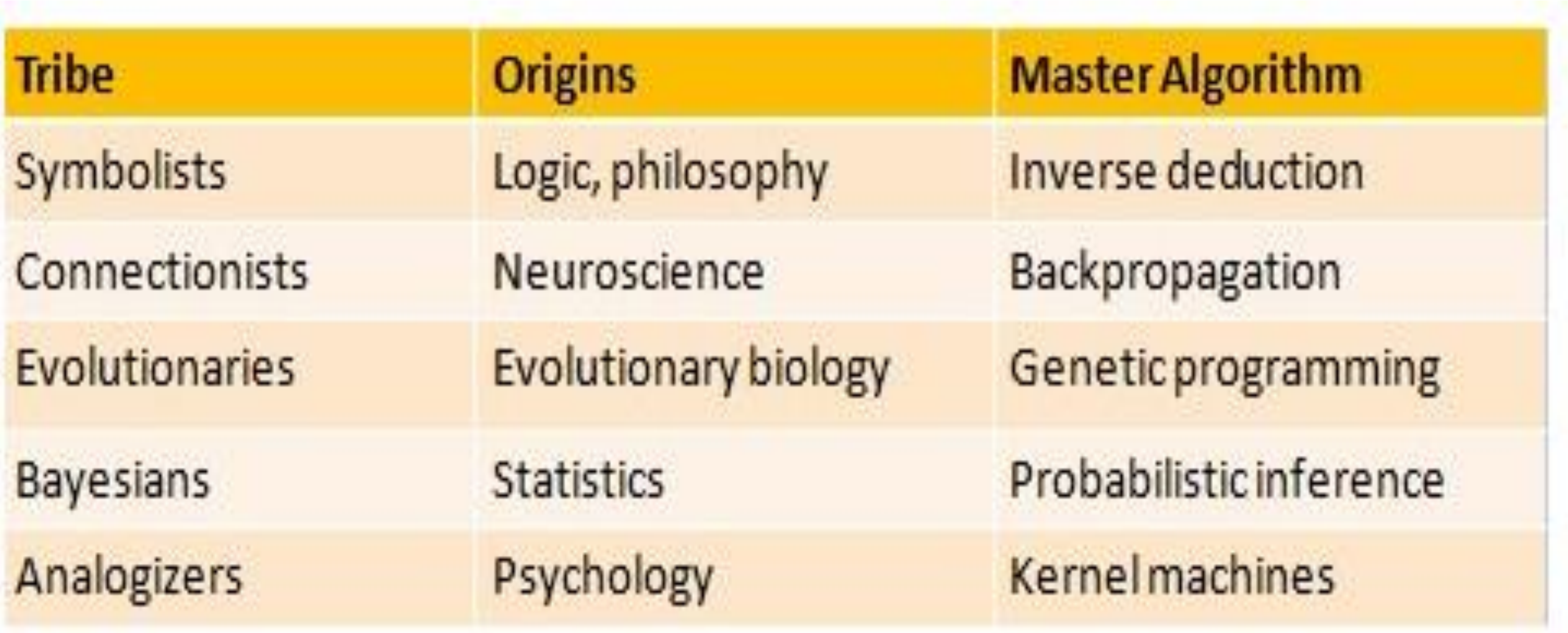
Optimize a function in light of constraints ("going as high as you can while staying on the road")

Favored algorithm

Support vectors

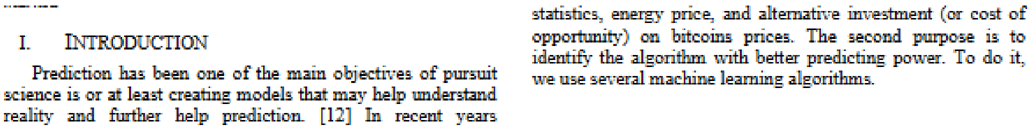
Source: Pedro Domingos, *The Master Algorithm*, 2015

Machine Learning



Tribe	Origins	Master Algorithm
Symbolists	Logic, philosophy	Inverse deduction
Connectionists	Neuroscience	Backpropagation
Evolutionaries	Evolutionary biology	Genetic programming
Bayesians	Statistics	Probabilistic inference
Analogizers	Psychology	Kernel machines

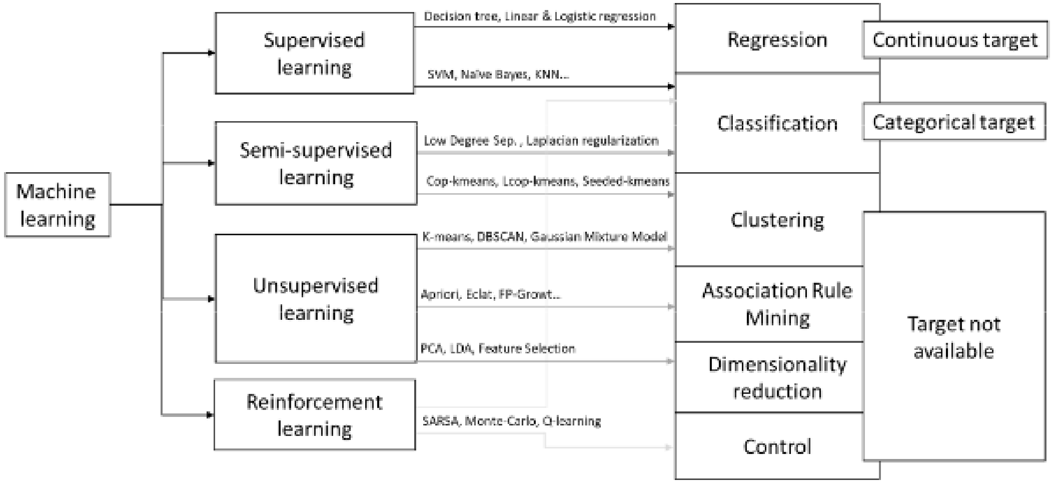
- Machine Learning Algorithms



I. INTRODUCTION

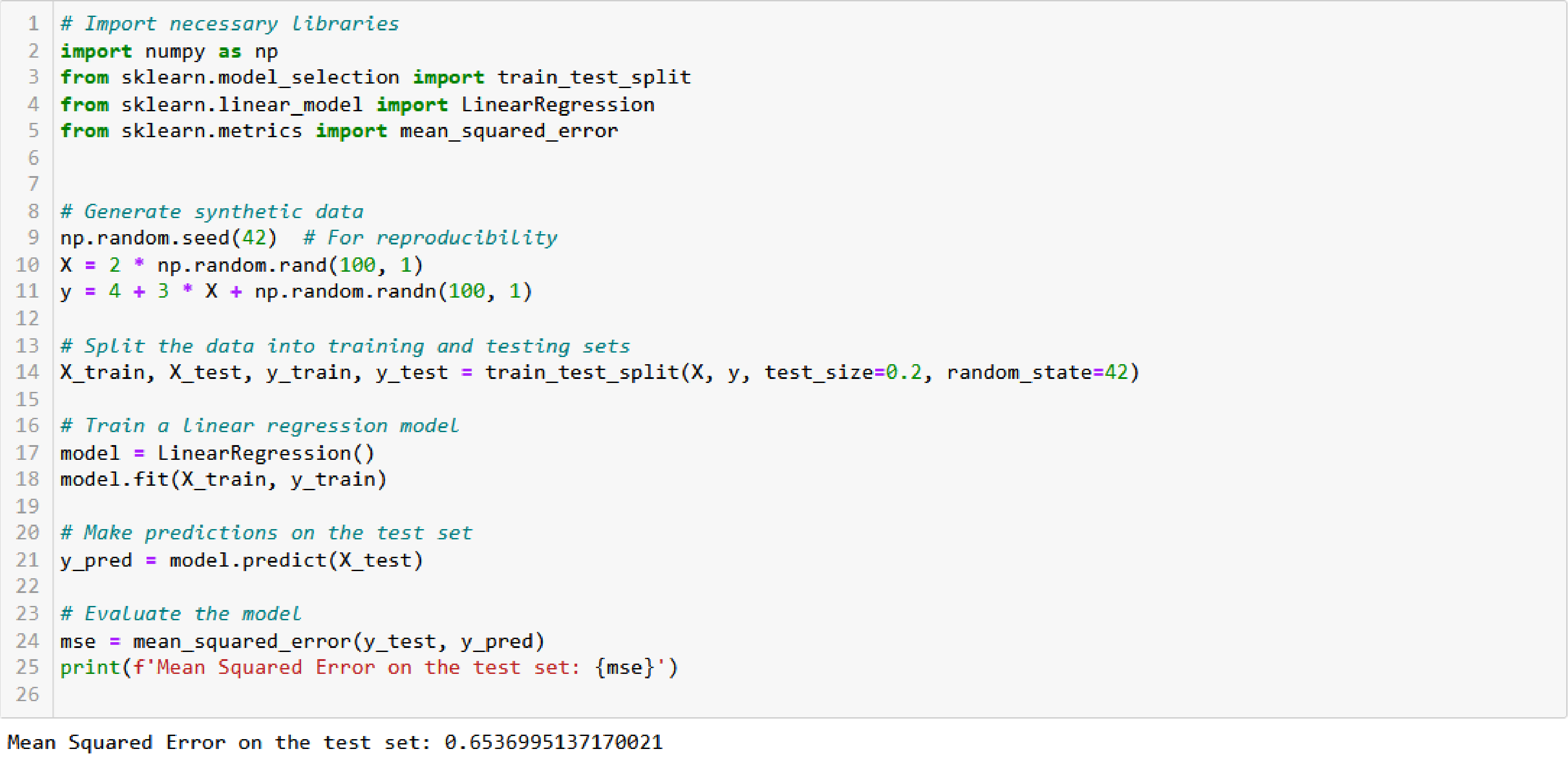
Prediction has been one of the main objectives of pursuit science is or at least creating models that may help understand reality and further help prediction. [12] In recent years

statistics, energy price, and alternative investment (or cost of opportunity) on bitcoins prices. The second purpose is to identify the algorithm with better predicting power. To do it, we use several machine learning algorithms.



- Aparicio, Romao & Costa (2022)

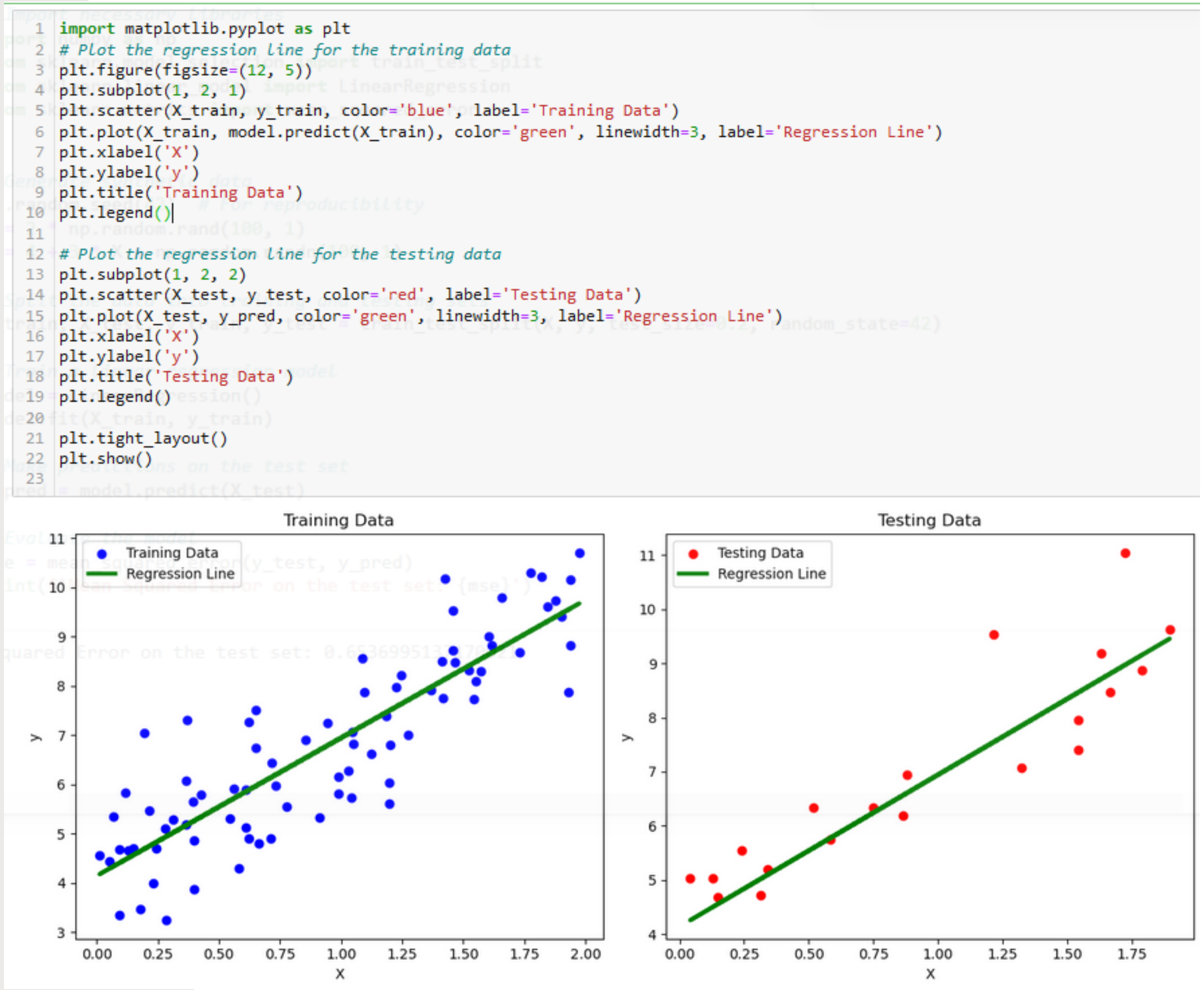
Example of supervised Model



```python
# Import necessary libraries
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error


# Generate synthetic data
np.random.seed(42)  # For reproducibility
X = 2 * np.random.rand(100, 1)
y = 4 + 3 * X + np.random.randn(100, 1)

# Split the data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# Train a linear regression model
model = LinearRegression()
model.fit(X_train, y_train)

# Make predictions on the test set
y_pred = model.predict(X_test)

# Evaluate the model
mse = mean_squared_error(y_test, y_pred)
print(f'Mean Squared Error on the test set: {mse}')

```

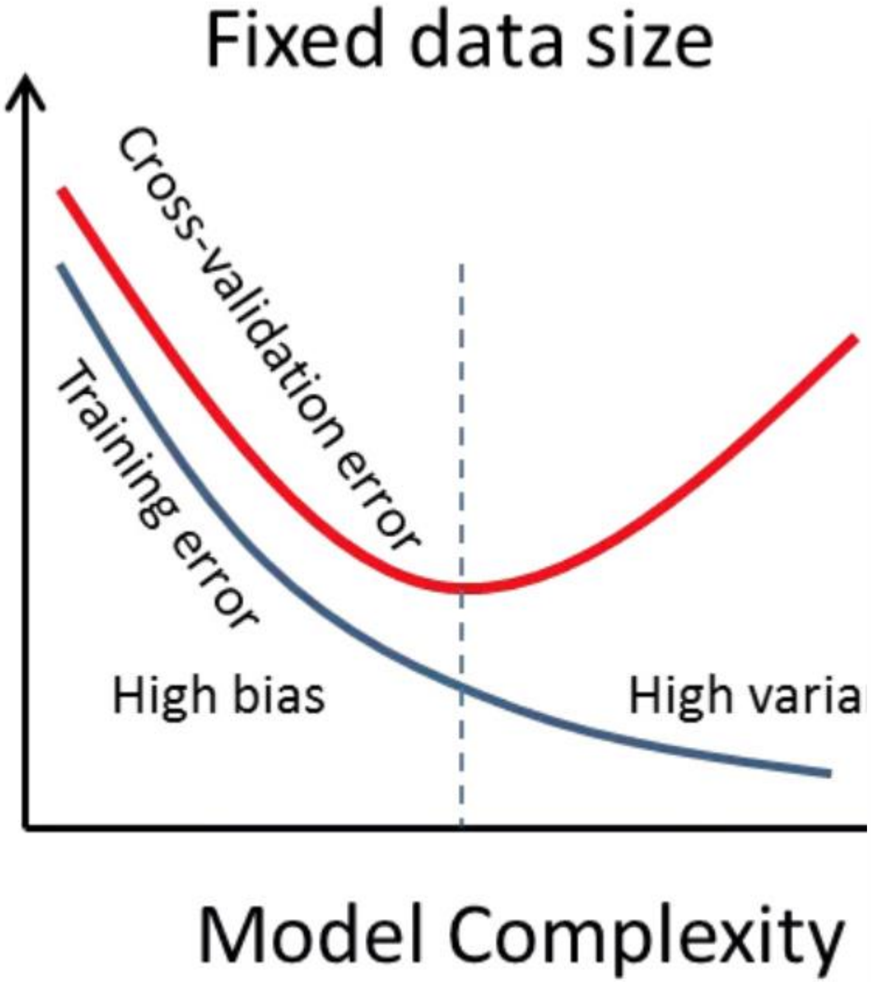
```
Mean Squared Error on the test set: 0.6536995137170021
```

Example of supervised Model



```python
import matplotlib.pyplot as plt
# Plot the regression line for the training data
plt.figure(figsize=(12, 5))
plt.subplot(1, 2, 1)
plt.scatter(X_train, y_train, color='blue', label='Training Data')
plt.plot(X_train, model.predict(X_train), color='green', linewidth=3, label='Regression Line')
plt.xlabel('X')
plt.ylabel('y')
plt.title('Training Data')
plt.legend()

# Plot the regression line for the testing data
plt.subplot(1, 2, 2)
plt.scatter(X_test, y_test, color='red', label='Testing Data')
plt.plot(X_test, y_pred, color='green', linewidth=3, label='Regression Line')
plt.xlabel('X')
plt.ylabel('y')
plt.title('Testing Data')
plt.legend()

plt.tight_layout()
plt.show()
```

Machine Learning



- **Train-Validate-Test**

- **Step 1: Making the model examine data.**
- **Step 2: Making the model learn from its mistakes.**
- **Step 3: Making a conclusion on how well the model performs**

	Prediction	Inference
Goal	Robust model using all predictors to accurately predict the outcome variable (Y) with high accuracy and low error.	Estimate the relationship between an outcome variable and predictor variable(s), while accounting for confounding factors.
Question Answered	How can I accurately predict new data points?	What do the relationships between the variables signify?
Example	Predicting house prices based on features like size, location, and number of bedrooms using regression models.	Inferring the impact of education level on income while controlling for factors such as experience and occupation using linear regression analysis.

SOMEONE SENT ME ANOTHER ANONYMOUS EMAIL WITH A LINK TO AN ARTICLE ABOUT THE WORLD'S WORST BOSSES.
Dilbert.com DilbertCartoonist@gmail.com
I GET ONE OF THOSE EMAILS EVERY TIME I LEAVE YOUR CUBICLE. DID YOU THINK I WOULDN'T NOTICE THE CORRELATION?
11-28-11 ©2011 Scott Adams, Inc./Dist. by Universal Uclick
CORRELATION DOES NOT IMPLY CAUSATION.
SEND

Inference

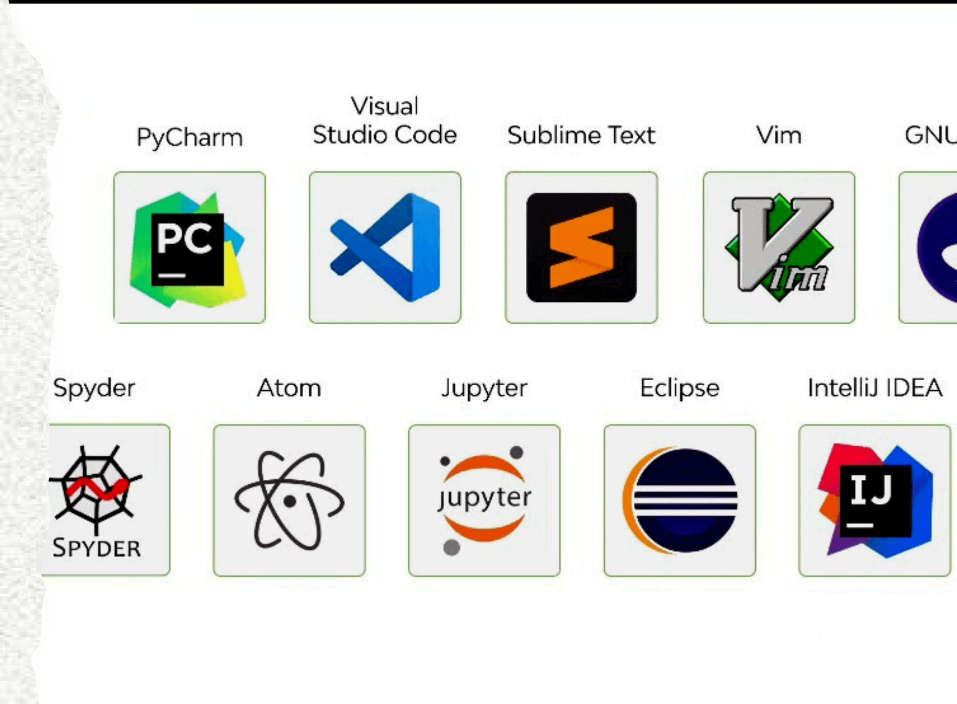
- Given a dataset, the purpose is to infer how the output is generated as a function of the data.
- Use the model to learn about the data generation process.
- Understand the way the independent variables X affect the target variable Y.
- Ex: find out what the effect of passenger gender, class and age, has on surviving the Titanic Disaster
- Model interpretability is a necessity for inference

Prediction

- Use the model to predict the outcomes for new data points.
- When performing predictions over data, the purpose is estimating f in y=f(x)
- The purpose is not understanding the exact form of the estimated function, as far as it can perform predictions quite accurately.
- To be able to predict what the responses are going to be to future input variables.
- Ex: predict prices of oil

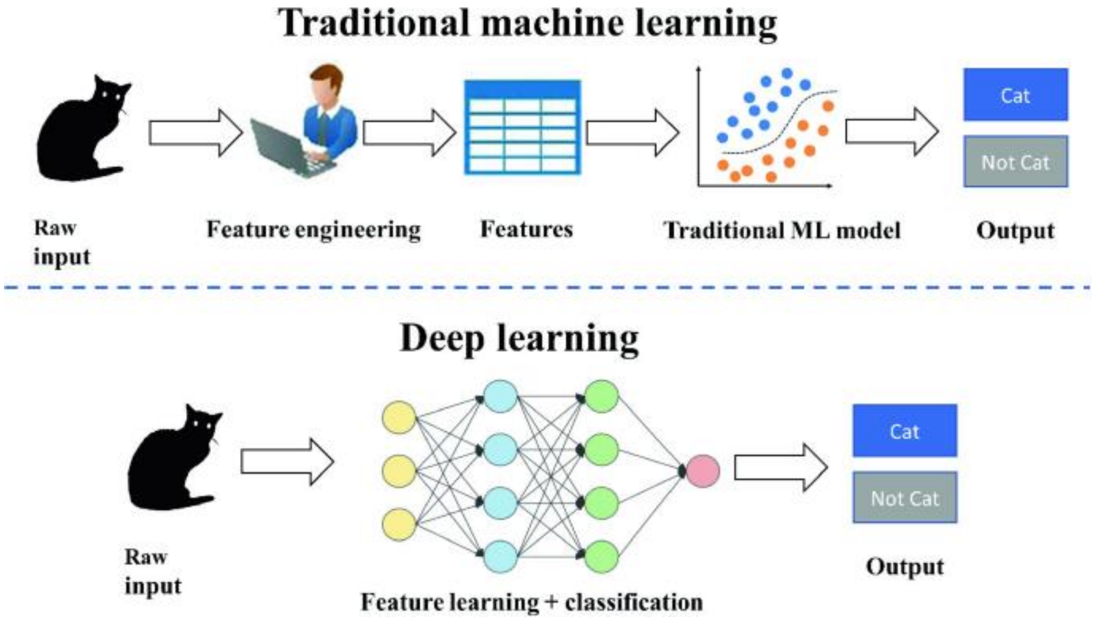
Machine Learning

- **Supervised Learning:**
  - Classification
  - Regression
- **Unsupervised Learning**
  - Clustering
  - Dimensional Reduction
- Reinforcement Learning



Deep learning

- is a subfield of machine learning
- focuses on the development and application of artificial neural networks, particularly deep neural networks.
    - composed of layers of interconnected nodes (artificial neurons) that can learn and make decisions.
- The term "deep" refers to the use of multiple layers in the neural network.



Natural Language Processing (NLP)

- subfield of artificial intelligence
- focuses on the interaction between computers and human language.
- The goal of NLP is to enable computers to understand, interpret, and generate human language in a way that is both meaningful and contextually relevant.
- involves the application of computational techniques and models to analyze and derive meaning from natural language data.
- Sentiment Analysis e increasing importance (Aparicio et al, 2021, Costa et al., 2021)

Emotion analysis of Portuguese Political Parties Communication over the covid-19 Pandemic

Joao Tiago Aparicio
Instituto Superior Técnico,
University of Lisbon
Lisbon, Portugal
joao.aparicio@tecnico.ulisboa.pt

João Salema de Sequeira
Instituto de Estudos Políticos
Universidade Católica Portuguesa
joaosalemasequeira@gmail.com

Carlos J. Costa
Instituto Universitário de Lisboa
(ISCTE-IUL), ISTAR-IUL,
Lisboa, Portugal
ISEG, Universidade de Lisboa,
Lisboa, Portugal
carlos.costa@acm.org

Abstract — In this paper, we explore the use of emotions in the Portuguese political parties' (with a seat in the Portuguese Parliament) communication as expressed by their official Twitter accounts, as of March 2020. The chosen period of our investigation is particularly interesting because political parties had a chance to communicate their views during a pandemic situation and over a period of one year. These views include possible solutions to face the crisis and their comments on the development of the whole situation. Using a standard lexicon we classified the amount of particular emotions in different tweets. Using this method we plotted the average positivity and negativity along time per party. We also analyzed the impact of each emotion to classify positivity using the present corpus. Finally, we considered some important words regarding the pandemic and their average positivity score. The analysis allows us to identify different approaches to participation in social media according to different strategies, more than political ideology.

Keywords - political communication; Portuguese political parties; Portuguese parliament; Portuguese; lexicon; sentiment analysis; emotions; visualization; social media; twitter; covid-19.

I. INTRODUCTION

Now-a-days, different political actors are increasingly using social media platforms to communicate their worldviews. American Presidents have used Twitter heavily to communicate their position in relation to specific ideas and to specific policies [5]. Hence it is essential to analyse what is being communicated and even more important how this communication is being done in order to best assess their impact. Political communication can help us explain the ups and downs of the electoral polls and the electoral success of a certain political party or individual in the following election.

The publication of *The Gutenberg Galaxy: The Making of Typographic Man* [8] considers the effects of social media in different human dimensions. However, a new empirical approach is needed, one that considers the effects of social networks or to put it simply a Zuckerberg Galaxy approach which demonstrates how Facebook, Twitter, and other social media are used and to what extent they have a more decisive influence on some of the voters, in comparison to the traditional media. In this context, the evolution in Natural Language Processing (NLP) and sentiment analysis is significant, however the political communication in Portugal has not yet been a subject of this kind of study, since the available models and lexicons are not yet adapted to European Portuguese. In this sense, we aim to answer the following question: What are the prevalent emotions in the Portuguese political parties' tweets during over the first year of the covid-19 pandemic?

The purpose of the work performed in this paper is to analyse the communication of the official Twitter accounts of the Portuguese political parties. The time frame ranges through 3200 last tweets, going as back as March 2020, when the first case of covid-19 was registered in the country. This period is specially interesting because political parties had a chance to communicate to the electorate their ideas in face of a social and economic crisis. It is important to take into consideration that the different parties tweeted with a different frequency, however the reality they were facing was one and the same.

II. LITERATURE REVIEW

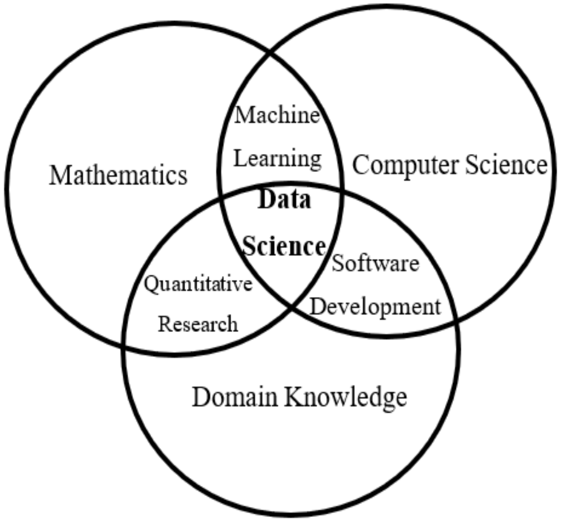
Sentiment analysis refers to using several approaches, such as: natural language processing, text analysis, computational linguistics, and biometrics, to systematically identify, extract, quantify, and study affective states and subjective information.

Emotions can be reactions to internal stimuli (such as thoughts or memories) or events in our environment. To analyze emotions, Mohammad and Turney [1] proposed a lexicon. This lexicon uses six emotions [2], [3]: joy, sadness, anger, fear, disgust, and surprise, along with how positive and negative the words are. These are a subset of the eight emotions proposed in Plutchik [4] which are still relevant today [10]. Recently the study of the impact of texts on such emotions has been done, namely in the USA political context [5]. This was done with a focus on awareness and topical emergence. However, there was no analysis over the emotion on the content of the message shared by the political parties, instead it was focused on its reception using Twitter users from states with opposing political views. This analysis was done over the covid-19 pandemic period, from 9th of March to the 13th of December, not encompassing any analysis over 2021. The study was composed of three stages, unigram frequencies identification, sentiment analysis and then topic modeling. The

2021 16th Iberian Conference on Information Systems and Technologies (CISTI)
23 – 26 June 2021, Chaves, Portugal
ISBN: 978-989-54659-1-0

Data Science

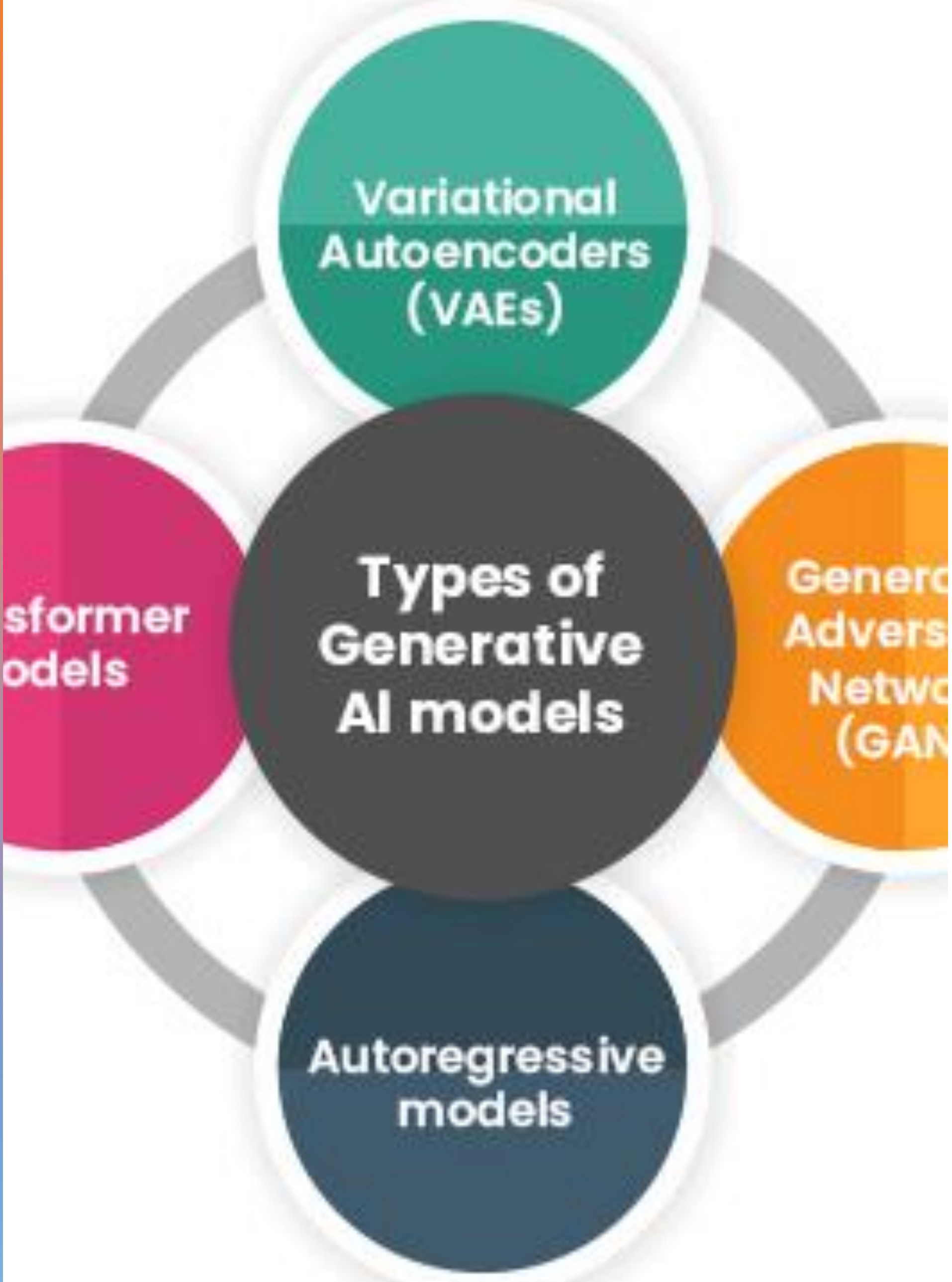
- includes techniques developed in some traditional fields like artificial intelligence, statistics or machine learning.



Aparicio et al. (2019).

Generative AI

- Class of AI algorithms and models that are designed to generate new, original content.
- Gen AI learn the underlying patterns and structures in the data and can generate novel outputs.
- *Instead of being trained on specific examples and then making predictions or classifications*
- These models are particularly good at creating content that resembles or is similar to the data they were trained on.

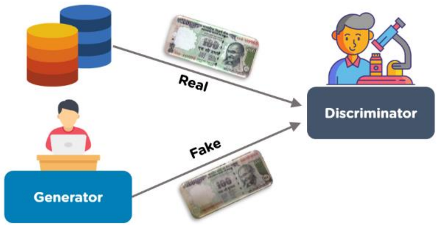

Variational Autoencoders (VAEs)
sformer odels
Types of Generative AI models
Genera Advers Netwo (GAN
Autoregressive models

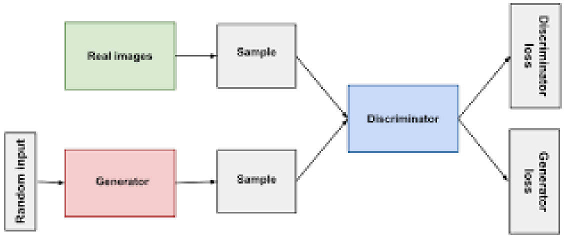
Types of generative AI models

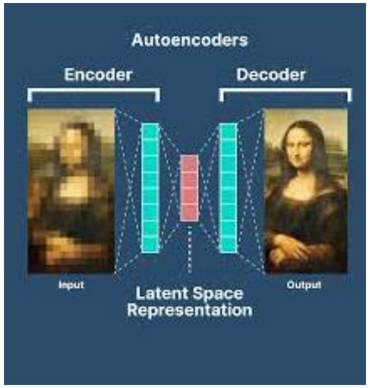
- Generative Adversarial Networks (GANs)
- Variational Autoencoders (VAEs)
- Autoregressive Models
- Recurrent Neural Networks (RNNs)
- Transformer-based Models
- Reinforcement Learning for Generative Tasks
- Generative AI for Data Privacy, Security and Governance.

Types of generative AI models

- Generative Adversarial Networks (GANs):
  - a generator and a discriminator are trained simultaneously through adversarial training.
- Variational Autoencoders (VAEs):
  - learn a probabilistic mapping from the observed data to a latent space.
  - Good to generate new samples from the learned latent space.
- Autoregressive Models:
  - the probability distribution of the next value in a sequence depends on the previous values.



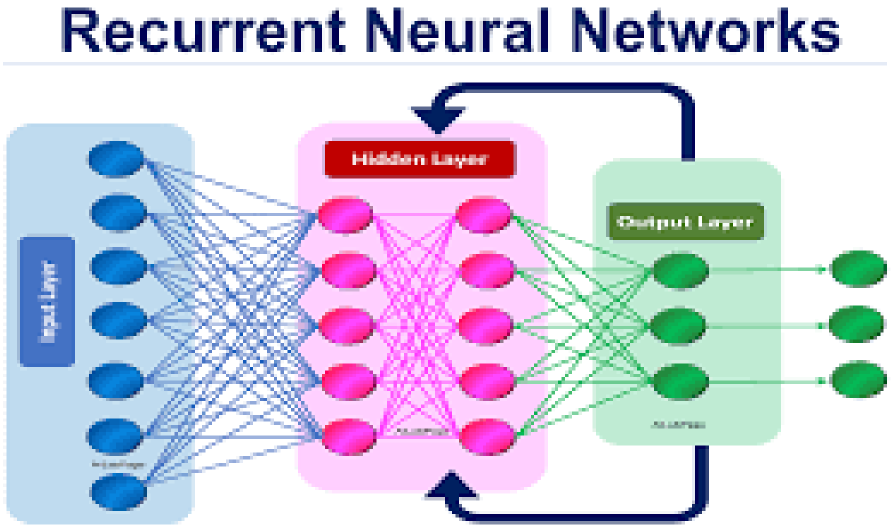




$$y_t = c + \sum_{i=1}^{p} a_{t-i} y_{y-i} + e_t$$

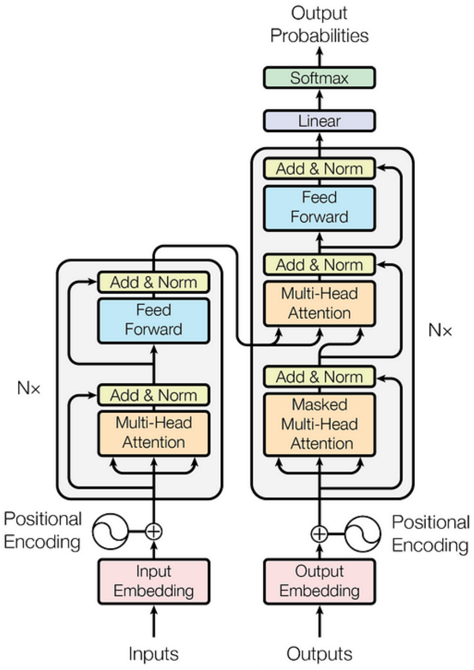
Types of generative AI models

- Recurrent Neural Networks (RNNs):
    - RNNs are commonly used for sequence tasks, including some generative tasks, they are not exclusively generative models.
    - Variants like LSTM and GRU are popular choices.
- Transformer-based Models:
    - Transformers, especially large language models.
- Reinforcement Learning for Generative Tasks:
    - can be used in conjunction with generative models, and this combination is powerful in scenarios where the generative model needs to produce sequences or structures guided by a reward signal.



BERT

Encoder

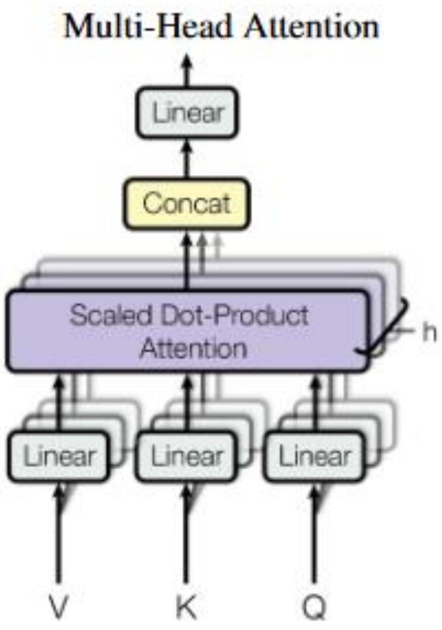


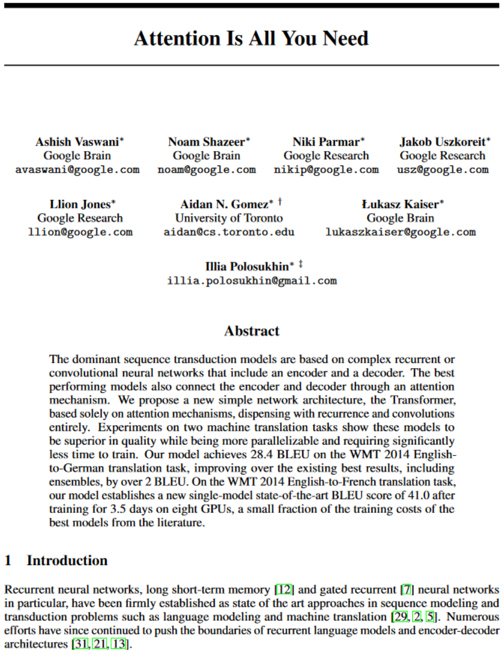
GPT

Decoder

Transformer

- Deep learning architecture based on the multi-head attention mechanism





Attention Is All You Need

Ashish Vaswani* Google Brain avaswani@google.com

Noam Shazeer* Google Brain noam@google.com

Niki Parmar* Google Research nikip@google.com

Jakob Uszkoreit* Google Research usz@google.com

Llion Jones* Google Research llion@google.com

Aidan N. Gomez* † University of Toronto aidan@cs.toronto.edu

Łukasz Kaiser* Google Brain lukaszkaiser@google.com

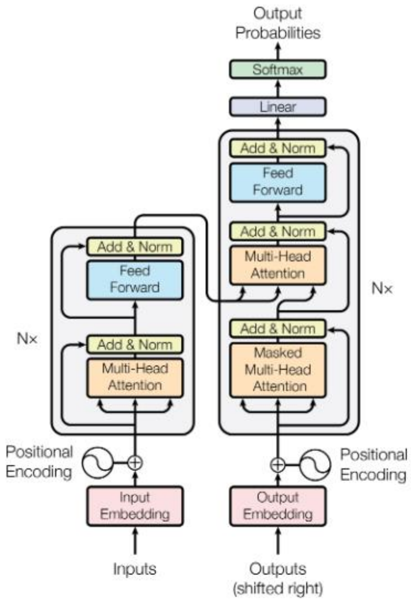
Illia Polosukhin* ‡ illia.polosukhin@gmail.com

Abstract

The dominant sequence transduction models are based on complex recurrent or convolutional neural networks that include an encoder and a decoder. The best performing models also connect the encoder and decoder through an attention mechanism. We propose a new simple network architecture, the Transformer, based solely on attention mechanisms, dispensing with recurrence and convolutions entirely. Experiments on two machine translation tasks show these models to be superior in quality while being more parallelizable and requiring significantly less time to train. Our model achieves 28.4 BLEU on the WMT 2014 English-to-German translation task, improving over the existing best results, including ensembles, by over 2 BLEU. On the WMT 2014 English-to-French translation task, our model establishes a new single-model state-of-the-art BLEU score of 41.0 after training for 3.5 days on eight GPUs, a small fraction of the training costs of the best models from the literature.

1 Introduction

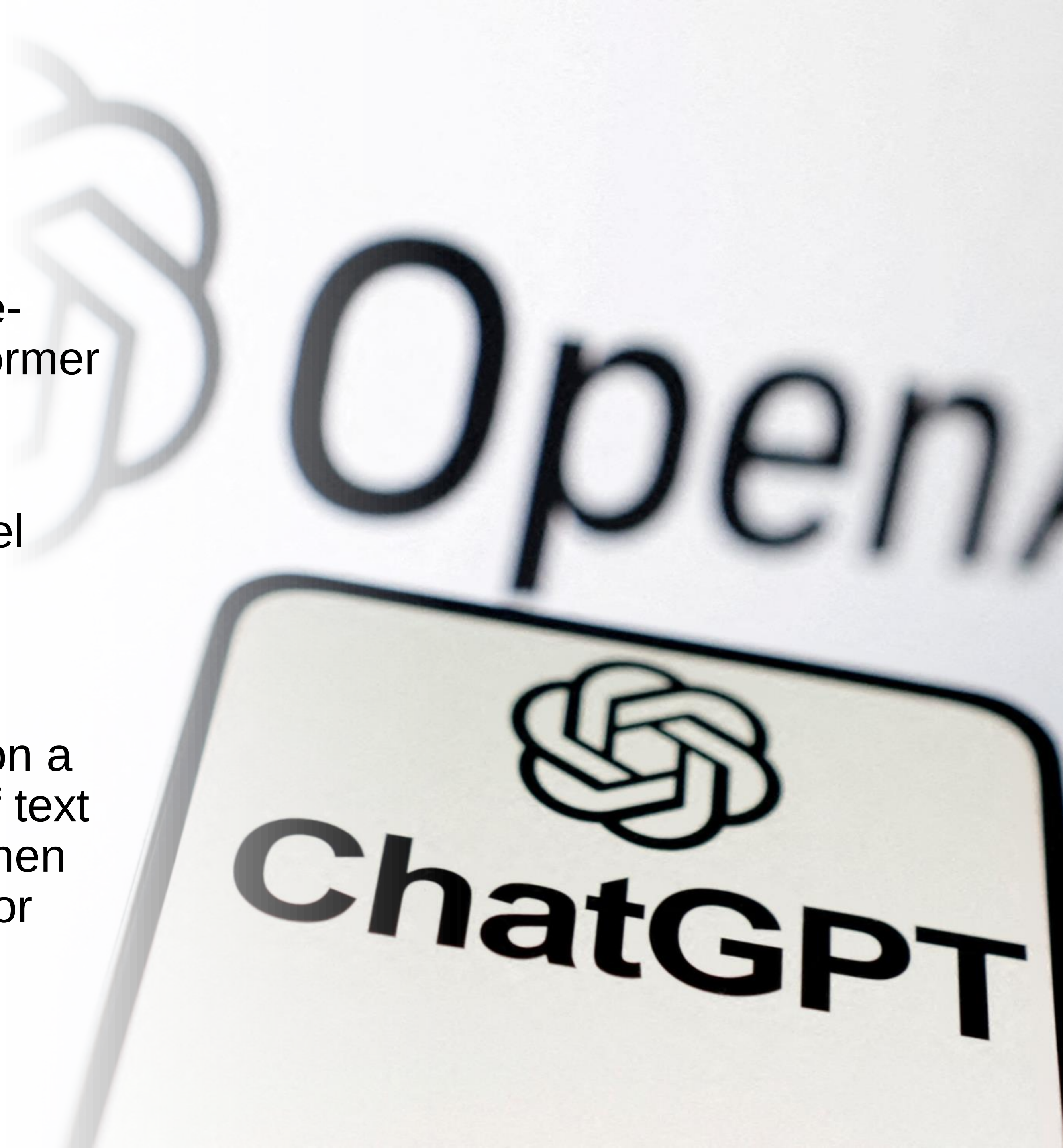
Recurrent neural networks, long short-term memory [12] and gated recurrent [7] neural networks in particular, have been firmly established as state of the art approaches in sequence modeling and transduction problems such as language modeling and machine translation [29, 2, 5]. Numerous efforts have since continued to push the boundaries of recurrent language models and encoder-decoder architectures [31, 21, 13].



Vaswani, et al. (2017)

GPT

- Generative Pre-trained Transformer
- Is a type of autoregressive language model that uses a transformer architecture.
- Is pre-trained on a large corpus of text data and can then be fine-tuned for specific tasks.



Google Gemini

Bard is a conversational AI chatbot powered by a combination of generative AI techniques, including:

- **Transformer-based models:**
  - Google's Pathways Language Model (PaLM) is used to generate text that is fluent, coherent, and grammatically correct.
- **Autoregressive models**
  - to predict the next word in a sequence, which helps to ensure that its responses are natural and engaging.
- **Reinforcement learning:**
  - it is rewarded for generating responses that are informative, comprehensive, and relevant to the user's query.

Feature	**LaMDA**	**PaLM**	**Gemini**
Release Date	2021	2022	December 2023
Focus	Conversational AI	General-purpose	Multimodal
Strengths	Realistic dialogue	Large & diverse dataset	Understanding & processing various data formats
Successor	Gemini/ PaLM	Gemini	N/A

Bibliography

Aparicio, J. T., de Sequeira, J. S., & Costa, C. J. (2021). Emotion analysis of portuguese political parties communication over the covid-19 pandemic. In 2021 16th Iberian Conference on Information Systems and Technologies (CISTI) (pp. 1-6). IEEE.

Aparicio, J. T., Romao, M., & Costa, C. J. (2022). Predicting Bitcoin prices: The effect of interest rate, search on the internet, and energy prices. In 2022 17th Iberian Conference on Information Systems and Technologies (CISTI) (pp. 1-5). IEEE.

Aparicio, S., Aparicio, J. T., & Costa, C. J. (2019,). Data Science and AI: trends analysis. In 2019 14th Iberian Conference on Information Systems and Technologies (CISTI) (pp. 1-6). IEEE.

Arriaga, A., & Costa, C. J. (2023, May). Modeling and Predicting Daily COVID-19 (SARS-CoV-2) Mortality in Portugal: The Impact of the Daily Cases, Vaccination, and Daily Temperatures. In Proceedings of International Conference on Information Technology and Applications: ICITA 2022 (pp. 275-285). Singapore: Springer Nature Singapore.

Costa, C. J., & Aparicio, J. T. (2020). POST-DS: A methodology to boost data science. In 2020 15th Iberian Conference on Information Systems and Technologies (CISTI) (pp. 1-6). IEEE.

Costa, C. J., & Aparicio, M. (2023). Applications of Data Science and Artificial Intelligence. Appl. Sci, 13, 9015.

Costa, C., Aparicio, M., & Aparicio, J. (2021). Sentiment analysis of portuguese political parties communication. In Proceedings of the 39th ACM International Conference on Design of Communication (pp. 63-69).

Custódio, J. P. G., Costa, C. J., & Carvalho, J. P. (2020). Success prediction of leads–A machine learning approach. In 2020 15th Iberian Conference on Information Systems and Technologies (CISTI) (pp. 1-6). IEEE.

Hajishirzi, R., & Costa, C. J. (2021). Artificial Intelligence as the core technology for the Digital Transformation process. In 2021 16th Iberian Conference on Information Systems and Technologies (CISTI) (pp. 1-6). IEEE.

Samadani, S., & Costa, C. J. (2021). Forecasting real estate prices in Portugal: A data science approach. In 2021 16th Iberian Conference on Information Systems and Technologies (CISTI) (pp. 1-6). IEEE.

Vaswani, A.; Shazeer, N.; Parmar, N. Uszkoreit, J.; Jones, L.; Gomez, A; Kaiser, Ł; Polosukhin, I (2017). "Attention is All you Need" Advances in Neural Information Processing Systems. Curran Associates, Inc. 30.