AI Concepts and Definitions

AI & SUSTAINABILITY IN VET EDUCATION

ERASMUS 2023-1-LT01-KA220-VET-000155506

Kristina Sutiene and Liepa Bikulciene

Kaunas University of Technology

Darica, Türkiye
Sep 16-20, 2024

Co-funded by the European Union

Outline

Table of Contents

Intro to AI

AI is a multi-use technology. Like electricity it can be applied in lots of different ways, to lots of different scenarios.

AI encompasses developing computer systems capable to exhibit intelligent behaviour.

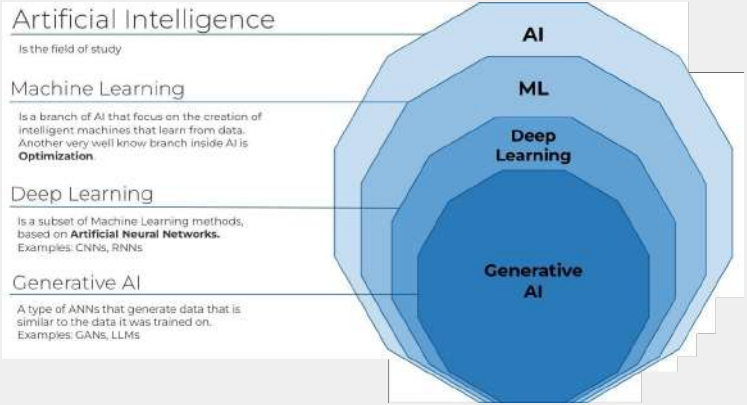
These tasks include:

- learning from experiences,
- interpreting complex data,
- making decisions based on the information gathered,
- understanding natural language,
- recognizing patterns or objects, etc.

The essence of AI lies in its ability to mimic cognitive functions associated with the human mind, such as learning and problem-solving.

Intro to AI

AI is the broadest concept, referring to any technique that enables machines to mimic human intelligence, encompassing reasoning, learning, and improving.



Core Components of AI

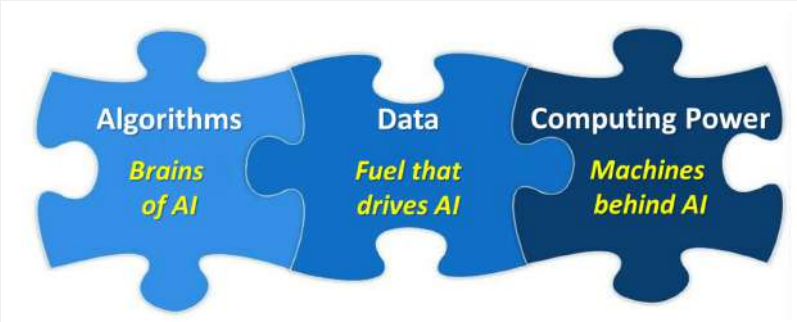
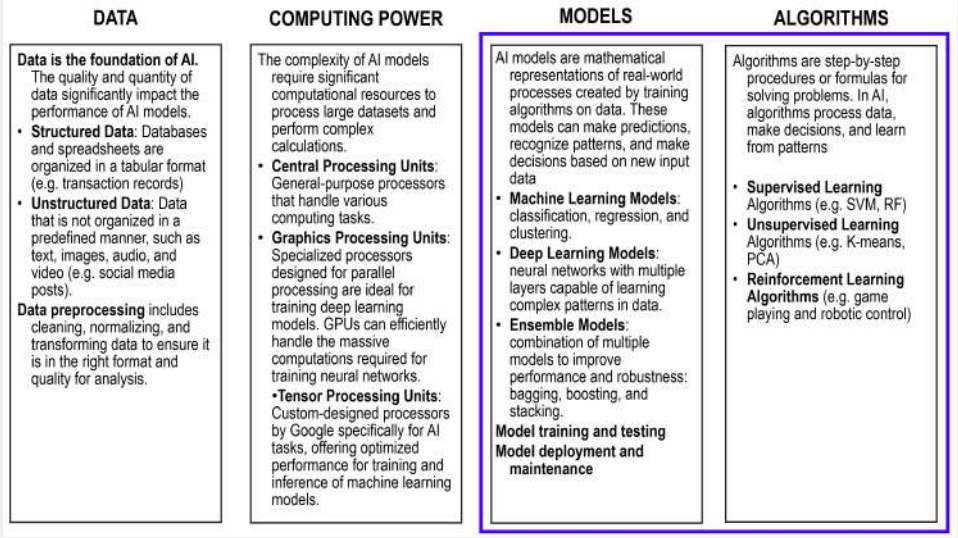
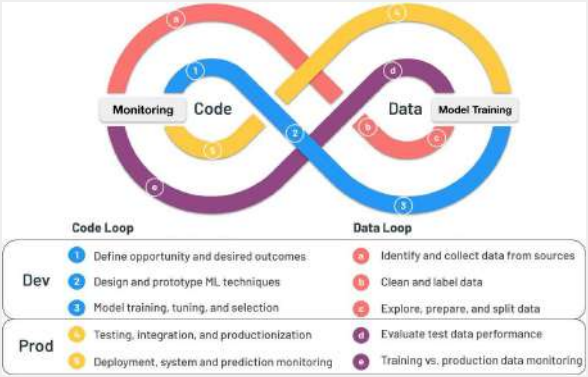


Figure: Algorithms tell computers what to do. Data tells computers what to learn. Computing power gives machines the power to learn and make decisions

Core Components of AI



Core Components of AI



A complex interplay of models, algorithms, and vast data processing capabilities.

[0] https://www.pachyderm.com/blog/what-is-a-data-pipeline-for-machine-learning/

AI concepts. Traditional programming vs AI

- 'Traditional programming' involves encoding human knowledge and experience into a set of precise rules that a computer can follow, step-by-step, which make the computer appear to respond intelligently.
- These rules, called algorithms, tell computers how to perform tasks and in traditional programming are often expressed in an 'IF-THEN-ELSE' format, which resembles a decision tree.
- For example to create a 'digital doctor', an algorithm might look like

```
IF the patient has a fever
    THEN prescribe Drug X
ELSE send the patient home
```

The intelligence in traditional computer systems comes directly from human knowledge and expertise being recorded into a format that a computer can process.

AI concepts. Traditional programming vs AI

- To develop a useful and reliable 'digital doctor', a huge number of rules and exceptions would be required that the system would very quickly become very large, very complicated, and unlikely to capture all of a real doctor's expertise and experience gained over time.
- Therefore traditional programming of 'intelligent systems' is best suited to constrained environments which do not change much over time and where the rules can be strictly defined.

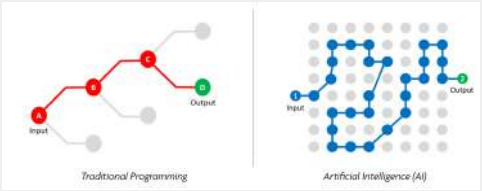


Figure: AI systems can also continuously 'learn from experience' and therefore not possible to program that logic by a series of defined steps.

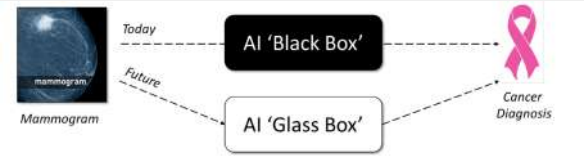
AI concepts. AI and XAI

AI systems can also continuously 'learn from experience' and therefore not possible to program that logic by a series of defined steps and decisions.

- An important concept in AI called 'Explainable AI (XAI)'.
- It is possible to verify the output of an AI system as correct (for example did a patient really have cancer or not), but it may not be possible to know exactly how the AI system got to a decision or outcome, as the AI has calculated the 'interim steps' by itself, without human programming of the specific steps to be taken.
- AI systems are therefore sometimes referred to as 'black boxes' with details of exactly how they work and operate - in some cases - not fully known, or understood.

AI concepts. AI and XAI

XAI is therefore an active and ongoing field in which researchers are trying to understand and articulate how AI systems reach an output. This is crucial to improve trust in AI systems (i.e. to understand the magic inside the black box and to turn a 'black box' into a 'glass box').



AI concepts. (Big) Data Concepts

"Data is the fuel" behind modern computing and AI algorithms, allowing them to learn, find relationships in data and make informed predictions and decisions.

Data	Bytes	Size
1 Kilobyte (KB)	1,000	10^3
1 Megabyte (MB)	1,000,000	10^6
1 Gigabyte (GB)	1,000,000,000	10^9
1 Terabyte (TB)	1,000,000,000,000	10^{12}
1 Petabyte (PB)	1,000,000,000,000,000	10^{15}
1 Exabyte (EB)	1,000,000,000,000,000,000	10^{18}
1 Zettabyte (ZB)	1,000,000,000,000,000,000,000	10^{21}

Figure: Paper 'Artificial Intelligence Threats & Opportunities' (EU Parlament, 2023) estimated that by 2025 the volume of data produced in the world each year could be 175 zettabytes

Data	Bytes	Size
1 Ronnabyte (RB)	1,000,000,000,000,000,000,000,000,000	10^{27}
1 Quettabyte (QB)	1,000,000,000,000,000,000,000,000,000,000	10^{30}

Figure: The volume of digital data in the world is rising so fast, that in 2022 scientists formally agreed two new units of measurement for the first time in 31 years: Ronnabyte (RB) and Quettabyte (QB)

AI concepts. (Big) Data Concepts

This data explosion is largely driven by three factors:

- Increasing numbers of smartphones and internet devices: the number of devices connected to the internet grows every year.
- Growth of social media: social media platforms generate many terabytes of text, video and audio data every day.
- New data collection and storage technologies: new technologies, such as sensors placed throughout our cities, energy grids and transport infrastructure, are generating huge amounts of data on human usage patterns. Similarly the 'Internet of Things (IoT)', which are the networked everyday items (such as fitness trackers and Smart TV's) are growing in popularity and producing new streams of data.

AI concepts. (Big) Data Concepts

- This increasing volume of data is also leading to a significant rise in cloud computing and data storage centres.
- Companies have been working towards becoming more data-driven for many years, with some now using cloud computing to store and manage their vast quantities of data on remote cloud-based servers (rather than on premises computers), in either 'data lakes' or 'data warehouses' (or both).

AI concepts. (Big) Data Concepts

- A data warehouse is a repository that stores structured data, in a highly organised and optimised way. This structured way of storage can be easily searched and is very good at quickly finding specific pieces of information.
- A data lake is a repository that can store all types of unstructured data, such as social media posts, images, videos, customer feedback surveys, or online reviews. This form of storage is good for collecting many different types of data and for keeping all the relevant information (within privacy limits), even when you're not sure how, or when, it could be used in the future.

When the best aspects of a data warehouse and data lake are merged into one data management solution it is called a data lakehouse.

AI concepts. (Big) Data Concepts

In particular 'big data' supports the rise of AI in the following ways :

- 1. Training Data: AI systems require substantial amounts of training data to learn patterns. This makes their predictions and decisions more accurate.
- 2. Feature Extraction: large and diverse datasets help AI algorithms to accurately identify and extract features from within data.
- 3. Model Performance: big data enhances the performance and robustness of AI models, allowing them to handle a wide range of situations and inputs.
- 4. Predictive Analytics: large datasets and AI can be used by businesses to forecast trends, regime shifts, reasons of changes, etc.

AI concepts. (Big) Data Concepts

In particular 'big data' supports the rise of AI in the following ways :

- 5. Personalisation: when AI systems can access large volumes of data on human behaviours, preferences, and interactions they can provide tailored suggestions and experiences for users.
- 6. Natural Language Processing (NLP): NLP models, such as AI powered chatbots, virtual assistants and language translation systems, benefit from diverse and extensive language data, such as books, articles and social media posts.
- 7. Real-Time Decision Making: as the 'velocity' of big data accelerates AI systems can process and analyse this data in (near) real time to provide instant recommendations, or decisions that require extra warnings.

AI concepts. Computing Power

- Training AI systems typically requires a lot of data. This data can be very large and very complex and needs to be processed very quickly, which requires a lot of computing power.
- For example, sophisticated AI chatbots (such as Google Bard, Microsoft Copilot) were trained on datasets that included hundreds of billions of words (for comparison the Bible contains fewer than one million words). To download this volume data on a typical home internet connection and then process it for AI training could take hundreds of years on a standard computer. This task therefore requires a special type of computing power.

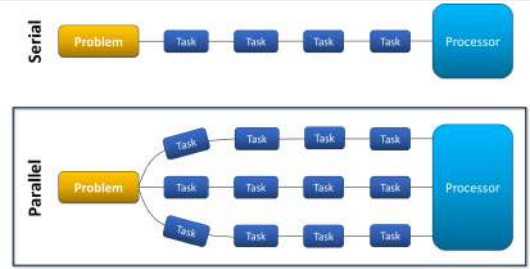
Computing power is crucial for training and deploying AI models. The complexity and size of modern AI models require significant computational resources to process large datasets and perform complex calculations.

AI concepts. Computing Power

Types of Computing Resources:

- Central Processing Units (CPUs): General-purpose processors that handle various computing tasks. While essential, they are often not sufficient for high-performance AI tasks.
- Graphics Processing Units (GPUs): Specialized processors designed for parallel processing are ideal for training deep learning models. GPUs can efficiently handle the massive computations required for training neural networks.
- Tensor Processing Units (TPUs): Custom-designed processors by Google specifically for AI tasks, offering optimized performance for training and inference of machine learning models.

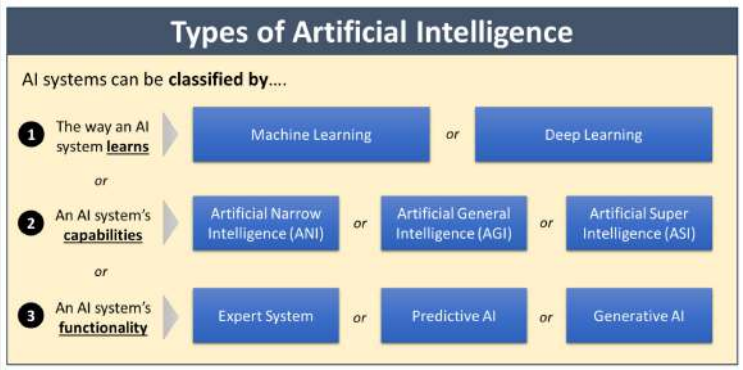
AI concepts. Computing Power



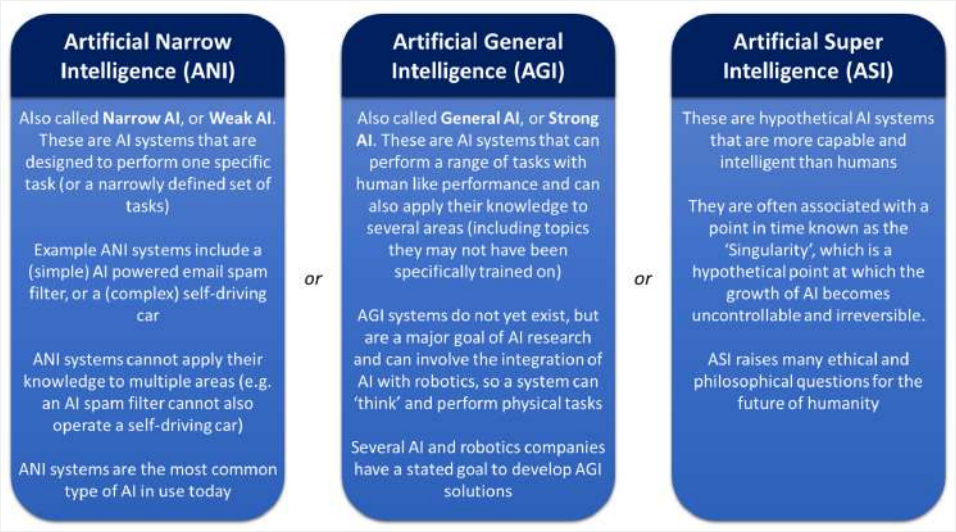
A key innovation is to process tasks in parallel (rather than sequentially as in normal computers), which turns out the accelerated parallel processing required to create lifelike graphics in computer games and movies is ideal for powering AI applications and especially deep learning systems, such as sophisticated computer vision aided assistants.

Types of AI

AI is a multi-use technology, with many different applications. There are therefore different ways that AI systems can be categorised, but they are normally segmented into one of three categories - either by the way they learn, by their capabilities or by their functionality.

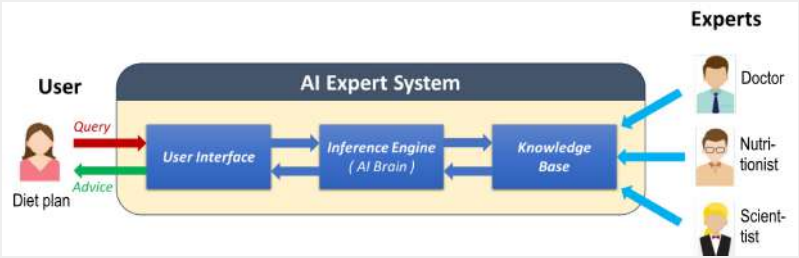


Types of AI


Artificial Narrow Intelligence (ANI)
Also called Narrow AI, or Weak AI. These are AI systems that are designed to perform one specific task (or a narrowly defined set of tasks)
Example ANI systems include a (simple) AI powered email spam filter, or a (complex) self-driving car
ANI systems cannot apply their knowledge to multiple areas (e.g. an AI spam filter cannot also operate a self-driving car)
ANI systems are the most common type of AI in use today
or
Artificial General Intelligence (AGI)
Also called General AI, or Strong AI. These are AI systems that can perform a range of tasks with human like performance and can also apply their knowledge to several areas (including topics they may not have been specifically trained on)
AGI systems do not yet exist, but are a major goal of AI research and can involve the integration of AI with robotics, so a system can 'think' and perform physical tasks
Several AI and robotics companies have a stated goal to develop AGI solutions
or
Artificial Super Intelligence (ASI)
These are hypothetical AI systems that are more capable and intelligent than humans
They are often associated with a point in time known as the 'Singularity', which is a hypothetical point at which the growth of AI becomes uncontrollable and irreversible.
ASI raises many ethical and philosophical questions for the future of humanity

Types of AI. Expert System

- An AI Expert System is a smart computer program that uses AI to simulate the expertise of humans in a specific area, such as for medical diagnosis, health monitoring or financial advice.
- Expert systems contain a knowledge base of rules and facts about a specific domain, which are combined with AI (in a system called an 'inference engine') to reason through problems and provide advice, or recommendations.



Types of AI. Predictive AI

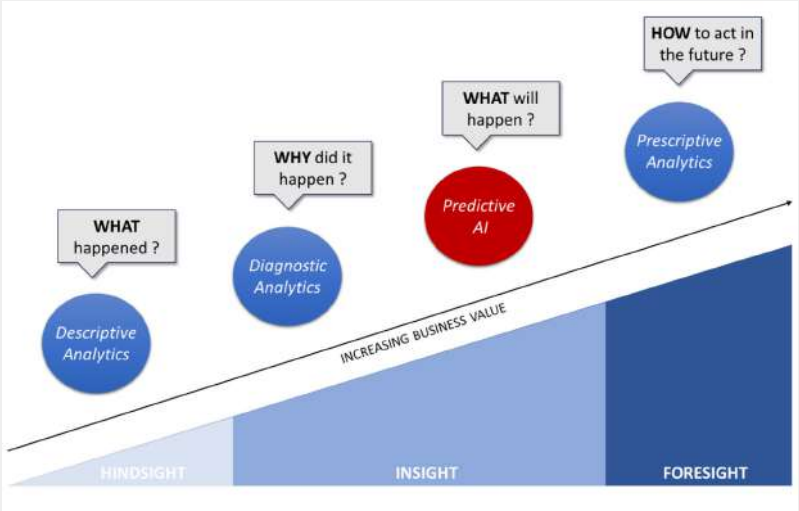
- Predictive AI is another way to classify an AI system's functionality and refers to AI that uses current and historical data to make predictions about future events, outcomes, or behaviours.
- They do this by analysing large amounts of data to identify patterns, trends and relationships that can help to anticipate the future.
- Predictive AI can also be combined with other forms of data analytics to create more comprehensive and powerful approaches to solving complex problems and gaining valuable business insights.

Types of AI. Predictive AI

Examples of other data analytical techniques that can be combined with predictive AI include:

- Descriptive Analytics: this summarises the past using historical data and helps to answer questions such as WHAT, WHERE and WHEN something happened by using methods such as dashboards, reports and other visual imagery to see and analyse patterns.
- Diagnostic Analytics: this aims to understand WHY something happened by investigating its root cause. These important insights from diagnostic analytics can also improve the accuracy of predictive AI models.
- Prescriptive Analytics: this suggests HOW to act in the future based on the data insights. By combining predictive AI forecasts with prescriptive recommendations it enables better planning and responses to potential future events.

Types of AI. Predictive AI



Types of AI. Generative AI

Generative AI is a relatively new form of AI that burst onto the global scene in late 2022 with the release of a sophisticated AI chatbot called ChatGPT.

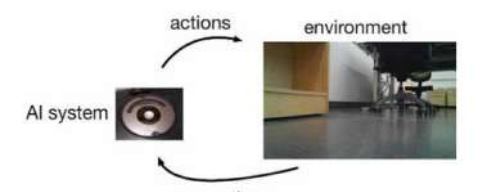
- AI chatbots are powered by generative AI, which is a type of AI technology that can produce new and creative content such as text, imagery, video and audio, from a user 'prompt'.
- It creates this new content based on the prompt and what it has learned from very large quantities of training data (sometimes many millions of internet pages) to produce a generative AI model.
- An example of this is foundation models which can be defined as "AI models trained on a broad data set that can applied to a wide range of downstream tasks.
- Generative AI has its origins in a 2017 Google research paper entitled "Attention Is All You Need", which introduced a new type of AI architecture for language understanding called 'Transformers'.

Table of Contents

AI and Machine Learning

ML theory and methods are instrumental for the analysis and design of AI.

- An AI system, typically referred to as an agent, interacts with its environment by executing (choosing between different) actions.
- The behaviour of an AI system is determined by how the perceptions made about the environment are used to form the next action. From an engineering point of view, AI aims at optimizing behaviour to maximize a long-term return. The optimization of behaviour is based solely on the perceptions made by the agent.



AI and Machine Learning

Let us consider some application domains:

- a personal diet assistant: perceived environment is the food preferences of the app user and their health condition; actions amount to personalized suggestions for healthy and tasty food; return is the increase in well-being or the reduction in public spending for health-care.
- a personal health assistant: perceptions given by current health condition (blood values, weight,...), lifestyle (preferred food, exercise plan); actions amount to personalized suggestions for changing lifestyle habits (less meat, more walking,...); return is measured via the level of well-being (or the reduction in public spending for health-care).

AI and Machine Learning

ML methods are used on different levels by an AI agent.

- On a lower level, ML methods help to extract the relevant information from raw data.
- ML methods are used to classify images into different categories which are then used an input for higher level functions of the artificial intelligence agent.
- To behave optimally, an agent is required to learn a good hypothesis for how its behaviour affects its environment. We can think of optimal behaviour as a consequent choice of actions that might be predicted by ML methods.

What sets AI applications apart from more traditional ML application is that there is an strong interaction between ML method and the data generation process.

Table of Contents

Data

Data is the collection of data points (instances). Data is typically considered as the most important component of any ML problem (and method) or AI system.

- We consider data as collections of individual data points which are atomic units of "information containers".
- Data points can represent text documents, signal samples of time series generated by sensors, entire time series generated by collections of sensors, frames within a single video, random variables, videos within a movie database, patients in the hospital, etc.

Data

- Tabular data: Structured data is stored in rows and columns, usually with each row corresponding to an observation and each column representing a variable in the dataset. An example is a spreadsheet of customer purchase histories.
- Text data: Unstructured textual data in natural language, code, or other formats. An example is a collection of posts and comments from an online forum.
- Image data: Digital representations of visual information that can train ML models to classify images, segment images, or perform other tasks. An example is a database of plant leaf images for identifying species of plants.

[0]The list is not exhaustive

Data

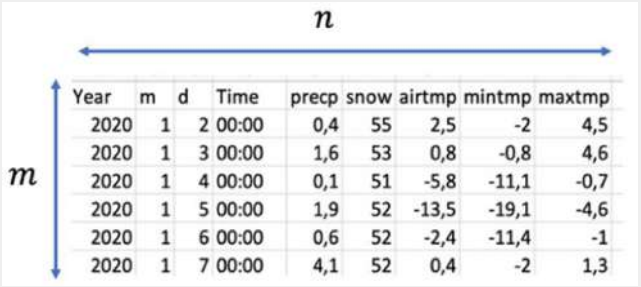
- Video data: A sequence of visual information over time that can train ML models to recognize actions, gestures, or objects in the footage. An example is a collection of sports videos for analyzing player movements.
- Audio data: Sound recordings, such as speech or music. An example is a set of voice recordings for training speech recognition models.
- Time-series data: Data collected over time that represents a sequence of observations or events. An example is historical stock price data.
- Graph data: Data representing a network or graph structure, such as social networks or road networks. An example is a graph that represents user connections in a social network.

[0]The list is not exhaustive

Data

- One practical requirement for a useful definition of data points is that we should have access to many of them. The model become more accurate for an increasing number of data points used for computing.
- A key parameter of a dataset is the number m of individual data points it contains. The number of data points within a dataset is also referred to as the sample size. Statistically, the larger the sample size m the better. However, there might be restrictions on computational resources (such as memory size) that limit the maximum sample size m that can be processed.

Data



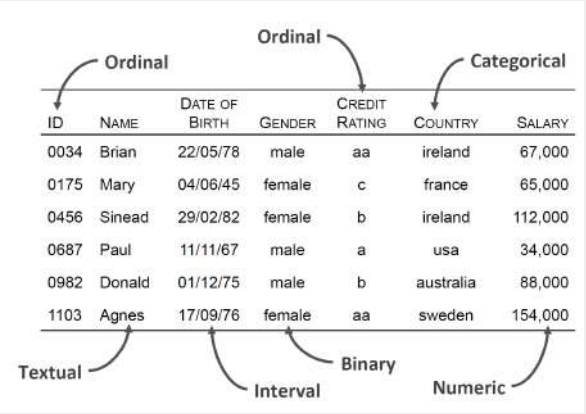
Year	m	d	Time	precp	snow	airtmp	mintmp	maxtmp
2020	1	2	00:00	0,4	55	2,5	-2	4,5
2020	1	3	00:00	1,6	53	0,8	-0,8	4,6
2020	1	4	00:00	0,1	51	-5,8	-11,1	-0,7
2020	1	5	00:00	1,9	52	-13,5	-19,1	-4,6
2020	1	6	00:00	0,6	52	-2,4	-11,4	-1
2020	1	7	00:00	4,1	52	0,4	-2	1,3

Figure: Two main parameters of a dataset are the number (sample size) m of individual data points that constitute the dataset and the number n of features used to characterize individual data points. The behaviour of ML methods typically depends crucially on the ratio m/n. The performance of ML methods typically improves with increasing m/n. As a rule of thumb, we should use datasets for which $m/n \gg 1$

Data

From the **computational view point** it is important to distinguish different types of data such as: Numeric: true numeric values that allow arithmetic operations (e.g., price) ; Interval: values that allow ordering and subtraction, but do not allow other arithmetic operations (e.g., time); Ordinal: values that allow ordering but do not permit arithmetic (e.g., small, medium, or large) ; Categorical: a finite set of values that cannot be ordered and allow no arithmetic (e.g., country); Binary: a set of just two values (e.g., gender); Textual: a free-form, usually short, text data (e.g., address).

Data



We may reduce this categorization to just two data types: continuous (numeric and interval types), and categorical (categorical, ordinal, binary, and textual types).

Data

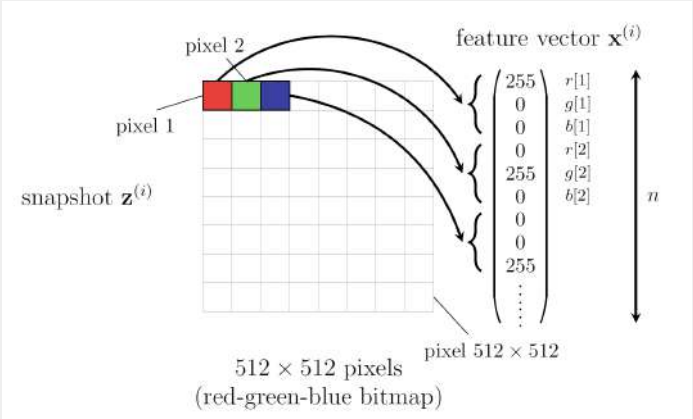


Figure: If the snapshot is stored as a 512×512 RGB bitmap, we could use as features the red-, green- and blue component of each pixel in the snapshot. The length of the feature vector would then be $n = 3 \cdot 512 \cdot 512 \approx 786000$

Data

From the **perspective of application domain and problem to be solved**: the data points could be classified into two different groups: features (or input variables, attributes) and labels (or target, output variable).

Typically, features are denoted by $X \in R^n$, while the label is denoted y (or in general $Y \in R^d$).

Data

The features X could be classified into: raw features or derived features.

- Raw features are features that come directly from raw data sources. For example, customer age, customer gender, loan amount, or insurance claim type are all descriptive features that we would most likely be able to transfer directly from a raw data source.
- Derived features do not exist in any raw data source, so they must be constructed from data in one or more raw data sources. For example, average customer purchases per month, loan-to-value ratios, or changes in usage frequencies for different periods are all descriptive features.

Data

There are a number of common derived feature types:

- Aggregates: These are aggregate measures defined over a group or period and are usually defined as the count, sum, average, minimum, or maximum of the values within a group.
- Flags: Flags are binary features that indicate presence or absence of some characteristic within a dataset.
- Ratios: Ratios are continuous features that capture the relationship between two or more raw data values.
- Mappings: Mappings are used to convert continuous features into categorical features and are often used to reduce the number of unique values that a model will have to deal with.

Data

The type of label y determines the problem from the modeling view point (loss function).

- Numeric Labels == Regression. Labels belong to a real number set. ML that aims at predicting a numeric label are referred to as regression method.
- Categorical Labels == Classification: label indicates the category or class to which data points belongs to. ML methods that aim at predicting such categorical labels are referred to as classification methods. For example, diagnosis of tumours as benign or malignant, or $0, 1$ or $-1, 1$ or yes, no.

Data

- We have a multi-class classification problem if data points belong to exactly one out of more than two categories (e.g., categories "red" vs. "green" and "blue"). If there are K different categories we might use the label values $1, 2, \ldots, K$.
- Ordinal Labels. Ordinal label values are somewhat in between numeric and categorical labels. Similar to categorical labels, ordinal labels take on values from a finite set. Moreover, similar to numeric labels, ordinal labels take on values from an ordered set. For example, categories "high risk" vs. "medium risk" and "low/no risk".

Data

- In the case of absence of any labeled data, ML methods can be useful for extracting relevant information or patterns from features X only.
- We refer to ML methods which do not require any labeled data points as unsupervised ML methods.

Data

- In the case of labeled data, we build supervised ML model by constructing a "good" predictor $h : X \to y$ which takes the features $x \in X$ of a data point as its input and outputs a predicted label (or output, or target) $\hat{y} = h(x) \in y$. A good predictor should be such that $\hat{y} \approx y$, i.e., the predicted label $\hat{y}$ is close (with small error $\hat{y} - y$ to the true underlying label y.

Data

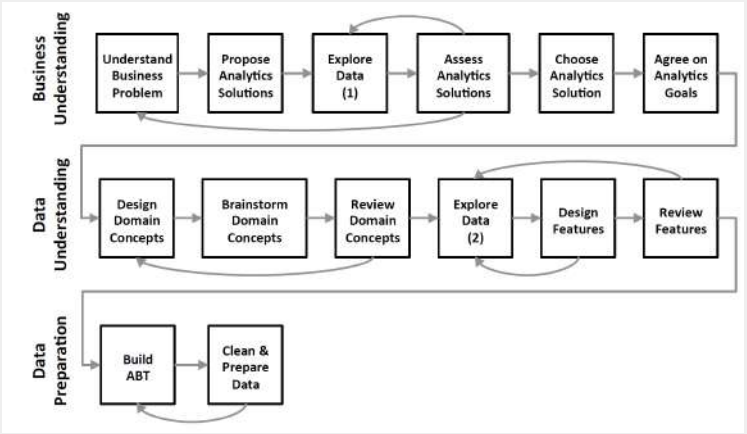


Figure: Tasks in the Business Understanding, Data Understanding, and Data Preparation phases

Table of Contents

Exploration Data Analysis (EDA)

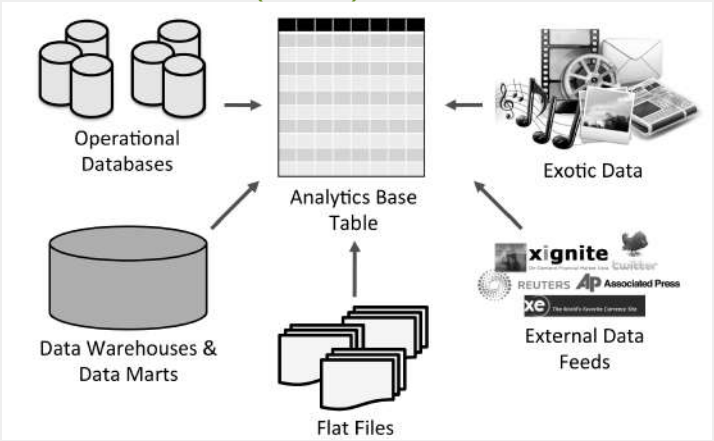


Figure: The different data sources typically combined to make a data set

Exploration Data Analysis (EDA)

There are two main goals in data exploration.

1. To fully understand the characteristics of the data. It is important that for each feature to understand characteristics such as the types of values a feature can take, the ranges into which the values in a feature fall, and how the values in a dataset for a feature are distributed across the range that they can take. We refer to this as getting to know the data.
2. To determine whether or not the data suffer from any data quality issue that could adversely affect the models that we build. Examples of typical data quality issues include an instance that is missing values for one or more descriptive features, an instance that has an extremely high value for a feature, or an instance that has an inappropriate level for a feature.

Exploration Data Analysis (EDA)

(a) Continuous Features

Feature	Count	% Miss.	Card.	Min.	1^{st} Qrt.	Mean	Median	3^{rd} Qrt.	Max.	Std. Dev.
—	—	—	—	—	—	—	—	—	—	—
—	—	—	—	—	—	—	—	—	—	—
—	—	—	—	—	—	—	—	—	—	—

(b) Categorical Features

Feature	Count	% Miss.	Card.	Mode	Mode Freq.	Mode %	2^{nd} Mode	2^{nd} Mode Freq.	2^{nd} Mode %
—	—	—	—	—	—	—	—	—	—
—	—	—	—	—	—	—	—	—	—
—	—	—	—	—	—	—	—	—	—

Figure: The structures of the tables included in a data quality report to describe (a) continuous features and (b) categorical features

Exploration Data Analysis (EDA)

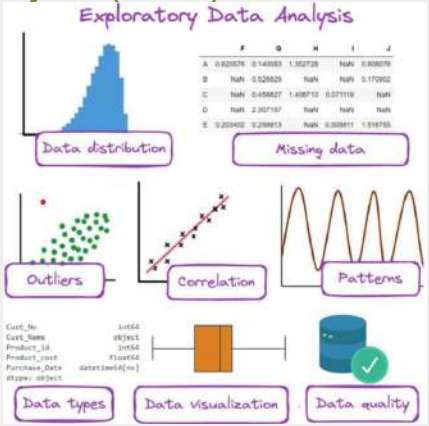


Figure: The Art of EDA - Techniques and Visualizations

[0] https://x.com/ingliguori/status/1742243587348209727

Exploration Data Analysis (EDA). Tools and techniques

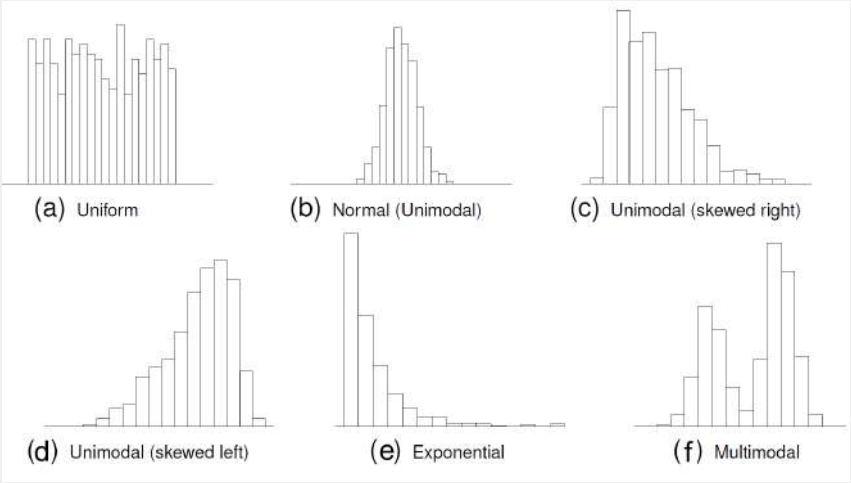
- Data cleaning addresses issues such as missing values, irrelevant information, irregular cardinality, and duplicate data.
- Outlier Detection is a crucial phase in exploratory data analysis. Using tools such as box plots or scatter plots can help recognize anomalies in your data.

Predictive modeling greatly benefits from a clean dataset, free of outliers. Specialized algorithms for outlier detection can be incorporated into your analysis process for further refinement and optimization.

Exploration Data Analysis (EDA). Tools and techniques

- Distribution analysis is a fundamental technique used in EDA. It helps us understand how the data points in our dataset are distributed across its range. It also helps in identifying any outliers and is often represented using histograms, density plots, and box plots.
- Addressing Skewness and Transformation: Addressing the issue of skewness is crucial in exploratory data analysis, especially in preparing data for predictive modeling. Transformations such as logarithmic or square root can help in normalizing skewed data.

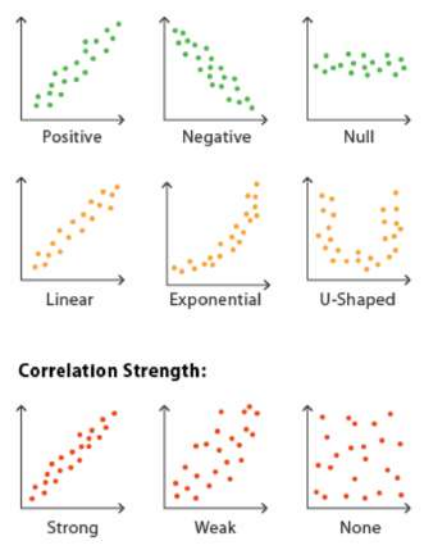
Exploration Data Analysis (EDA). Tools and techniques



Exploration Data Analysis (EDA). Tools and techniques

- Correlation Analysis helps to identify the relationships between various variables in our dataset. Understanding these dependencies is an important concept in machine learning, as it allows the selection of relevant predictors for your model and avoids multicollinearity.
- To deal with multicollinearity in predictive modeling, one popular method includes using the Variance Inflation Factor (VIF) to detect the severity of multicollinearity. You might also consider applying dimensionality reduction techniques such as Principal Component Analysis (PCA) or using regularization methods such as Ridge Regression or Lasso, which can handle multicollinearity effectively.

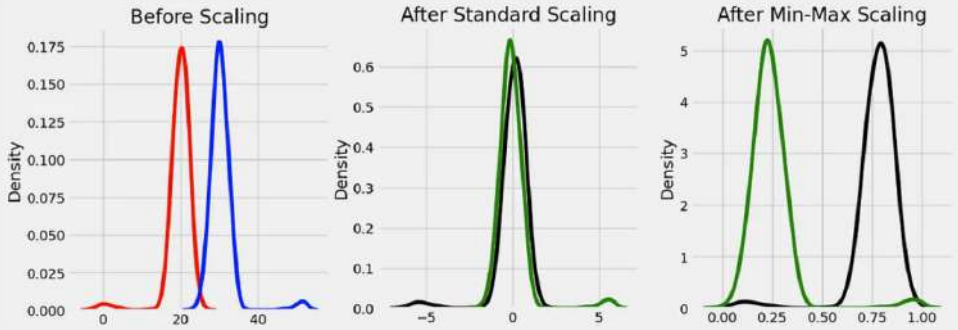
Exploration Data Analysis (EDA). Tools and techniques



Exploration Data Analysis (EDA). Tools and techniques

- Data Scaling (standardization) and Normalization are essential strategies for predictive modeling. The normalization is the process of converting data values to a common scale, while standardization is the process of converting data values to have the mean=0, and standard deviation=1, which ensures no single attribute dominates and makes easier interpretion.

Exploration Data Analysis (EDA). Tools and techniques



Exploration Data Analysis (EDA). Tools and techniques

- Handling Categorical Variables: Managing categorical variables is an essential step. It involves converting non-numeric data into a format that machine learning algorithms can understand. This is done through methods such as one-hot encoding or label encoding, aiding in better data processing.
- When you face scenarios with imbalanced datasets, it's critical to adopt strategies to balance your data for effective predictive modeling. This could involve oversampling the minority class, undersampling the majority class, or synthesizing new minority classes.

Exploration Data Analysis (EDA). Tools and techniques

- Time Series Visualization: Decoding patterns and seasonal trends through time series visualization is a powerful component.
- Preprocessing of time series data involves careful handling of gaps and shifts, understanding timestamps, and making necessary conversions. It's important to handle seasonality and trend components appropriately in time series data. These crucial steps ensure the data is ready and reliable for effective predictive modeling.

Exploration Data Analysis (EDA). Tools and techniques

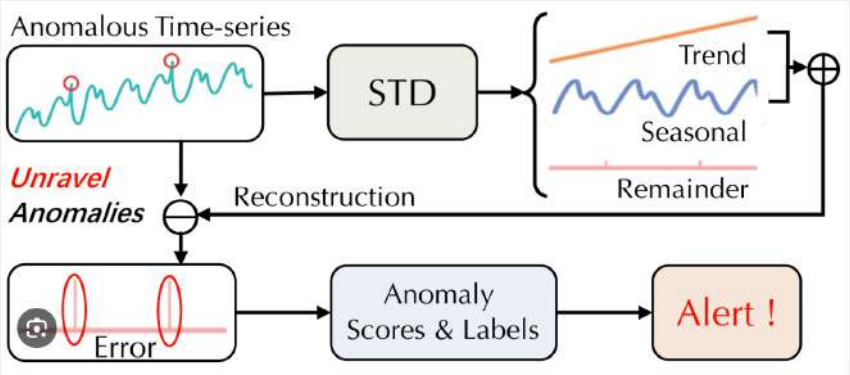


Figure: Preprocessing of time series data: seasonal-trend decomposition (STD) and anomaly detection

[0] https://arxiv.org/html/2310.00268v2

Exploration Data Analysis (EDA)

- Working with real-world datasets presents numerous challenges. Data analysts often encounter difficulties like insufficient data, noisy or high-dimensional data, imbalanced datasets, and tight deadlines. These are just some of the obstacles faced during EDA.
- However, utilizing the appropriate tools and implementing effective strategies can help overcome these challenges.
- The data preprocessing steps can greatly enhance the results of the models by improving the data's accuracy, thus adding value to the modeling process itself.

References

- Julia Simpson (2024) Introduction to Artificial Intelligence (AI) Technology https://cdn-dynmedia-1.microsoft.com/is/content/microsoftcorp/microsoft/final/en-us/microsoft-brand/documents/2024-wttc-introduction-to-ai.pdf
- Fredrik Filipsson (2024) Artificial Intelligence: An Introduction to AI Fundamentals https://redresscompliance.com/artificial-intelligence-an-introduction-to-ai-fundamentals/
- Dan Hendrycks (2024) Introduction to AI Safety, Ethics and Society https://www.aisafetybook.com/textbook
- MarkovML (2024) Exploratory Data Analysis in Predictive Modeling: Techniques & Strategies https://www.markovml.com/blog/exploratory-data-analysis
- GeeksforGeeks https://www.geeksforgeeks.org/
- Alexander Jung (2022) Machine Learning: The Basics https://alexjungaalto.github.io/MLBasicsBook.pdf
- John D. Kelleher, Brian Mac Namee and Aoife D'Arcy (2015) Fundamentals of Machine Learning for Predictive Data Analytics https://mitpress.mit.edu/9780262029445/fundamentals-of-machine-learning-for-predictive-data-analytics/

AI Concepts and Definitions

AI & SUSTAINABILITY IN VET EDUCATION

ERASMUS 2023-1-LT01-KA220-VET-000155506

Kristina Sutiene and Liepa Bikulciene

Kaunas University of Technology

Darica, Türkiye
Sep 16-20, 2024