Introduction to Machine Learning

CS-309

Sourav Kumar Dandapat

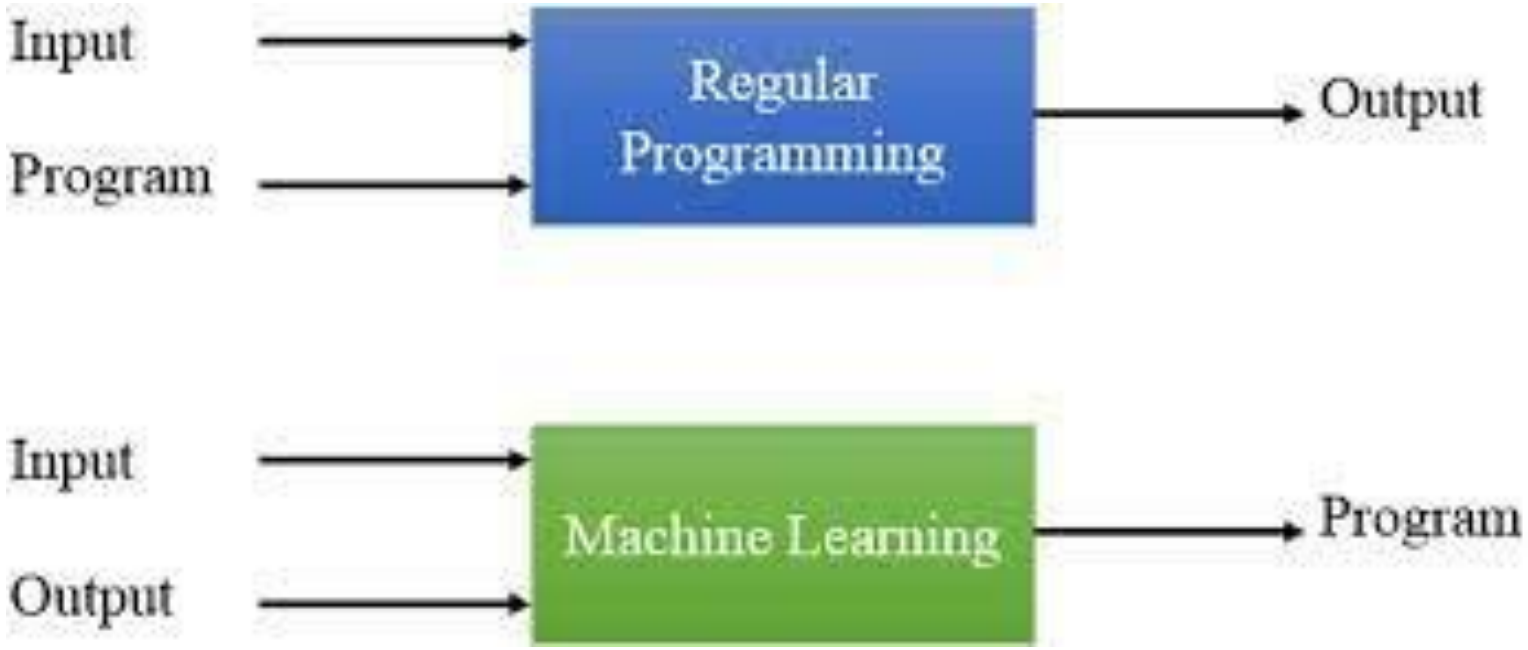
Text Books

- Pattern Recognition and Machine Learning by Christopher M. Bishop
- The elements of statistical learning by Hastie, Tibshirani and Friedman
- Machine Learning by Tom M. Mitchell

Evaluation Policy

- Quiz (4 quizzes will be taken out of which best 2 quiz marks will be considered, no extra quiz for missed quiz) : 25
- Midsem: 25
- Endsem:50

How it is different from traditional programming

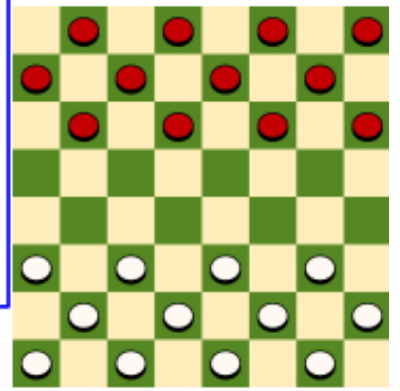


Definition of Machine Learning

Arthur Samuel (1959): Machine Learning is the field of study that gives the computer the ability to learn without being explicitly programmed.

A. L. Samuel*

Some Studies in Machine Learning
Using the Game of Checkers. II—Recent Progress



Photos from Wikipedia

Cont..

- Tom Mitchell (1998): 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.
- Experience (data): games played by the program (with itself)
- Performance measure: winning rate

History of Machine Learning

- **1950s - The Birth of Machine Learning:** The term "machine learning" was coined by Arthur Samuel in 1959.

- **1960s - Concept of Neural Networks:** The concept of artificial neural networks was introduced by Frank Rosenblatt. He created the perceptron and demonstrated its capabilities in image recognition tasks.

- **1970s - Decision Trees and Rule-Based Systems:** Decision tree algorithms, such as ID3 were introduced in the 1970s by Ross Quinlan. Rule-based systems, like MYCIN, demonstrated the use of expert systems in medical diagnosis.

- **1980s - Backpropagation and Support Vector Machines:** The backpropagation algorithm, a method for training artificial neural networks, was proposed in the 1980s. Support Vector Machines (SVMs) were also introduced in the 1980s.

- **1990s - Ensemble Learning and Boosting:** Ensemble learning techniques, such as bagging and boosting, gained popularity in the 1990s. Algorithms like Random Forests and AdaBoost demonstrated improved performance by combining multiple models.

- **2000s - Rise of Deep Learning:** CNN introduced.

- **2010s - Deep Learning Dominance:** The 2010s saw the dominance of deep learning across various domains, including computer vision, natural language processing, and speech recognition.

- **Present: Advances in Natural Language Processing and Reinforcement Learning** Recent years have seen remarkable progress in natural language processing, with models like OpenAI's GPT-3 demonstrating language understanding at unprecedented levels. Reinforcement learning has also made strides, with applications in robotics and game playing.

Types of learning

- **Supervised learning**
  - Learning mapping between input x and desired output y
  - Teacher gives me y's for the learning purposes
- **Unsupervised learning**
  - Learning relations between data components
  - No specific outputs given by a teacher
- **Reinforcement learning**
  - Learning mapping between input x and desired output y
  - Critic does not give me y's but instead a signal (reinforcement) of how good my answer was
- **Other types of learning:**
  - **Semi-Supervised Learning, Transfer Learning, Ensemble Learning**

Types of Learning

- **Supervised Learning:**
  - In supervised learning, the algorithm is trained on a labeled dataset, where the input data is paired with corresponding output labels.
  - **Objective:** The goal is to learn a mapping or relationship between the input variables (features) and the target output variable.
  - **Examples:** Classification tasks (e.g., spam detection, image recognition) and regression tasks (e.g., predicting house prices).

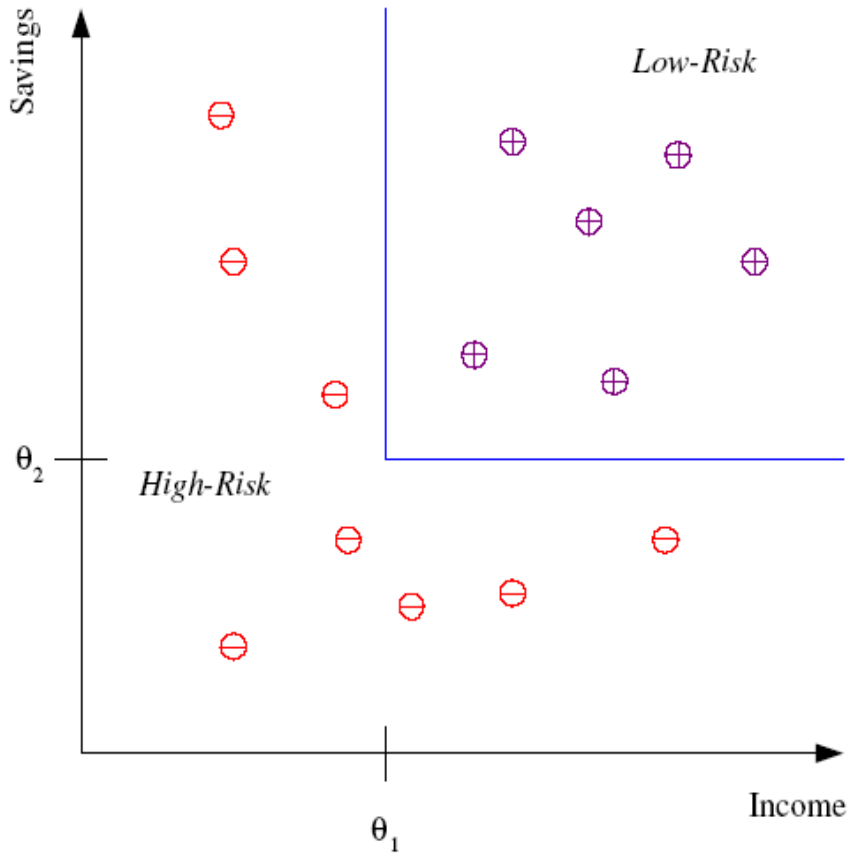
Types of Supervised Learning

- Supervised learning can be further categorized depending on the nature of the output variable and the specific goal of the model.
- **Classification:**
  - **In classification tasks, the goal is to predict the categorical class labels of new instances based on past observations.**
  - **Example: Spam detection (binary classification: spam or not spam), image classification (multiclass classification: recognizing different objects in images).**
- **Regression:**
  - **Regression involves predicting a continuous numerical output or value rather than discrete categories.**
  - **Example: Predicting house prices based on features like square footage, number of bedrooms, etc.**

Classification Example Credit scoring

- Differentiating between low-risk and high-risk customers from their *income* and *savings*

Discriminant: IF *income* > θ_1 AND *savings* > θ_2 THEN low-risk ELSE high-risk



Classification: Applications

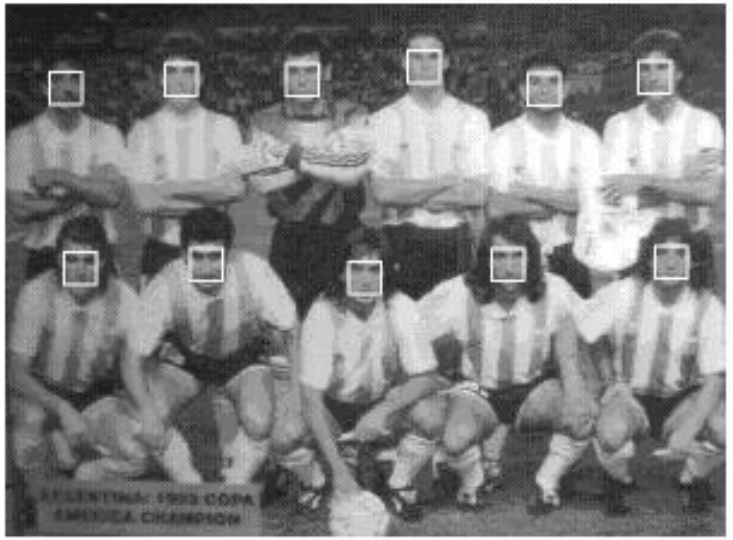
- Aka Pattern recognition
- **Face recognition:** **Pose, lighting, occlusion (glasses, beard), make-up, hair style**
- **Character recognition:** **Different handwriting styles.**
- **Speech recognition:**
  - **Use of a dictionary or the syntax of the language.**
  - **Sensor fusion: Combine multiple modalities; eg, visual (lip image) and acoustic for speech**
- **Medical diagnosis:** **From symptoms to illnesses**
- …

Face Recognition

Training examples of a person

Test images



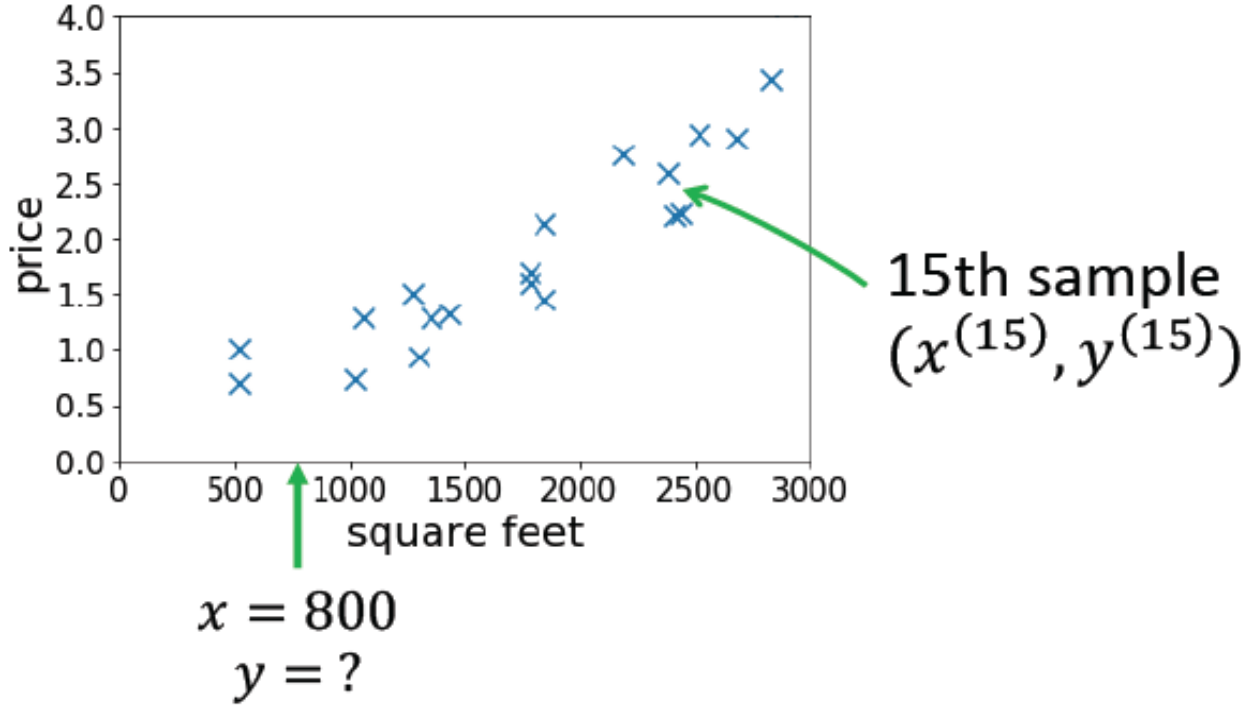







Regression Example: Price prediction for Housing

- Given: a dataset that contains n samples
- $(x^{(1)}, y^{(1)}),\ldots(x^{(n)}, y^{(n)})$
- Task: if a residence has x square feet, predict its price?



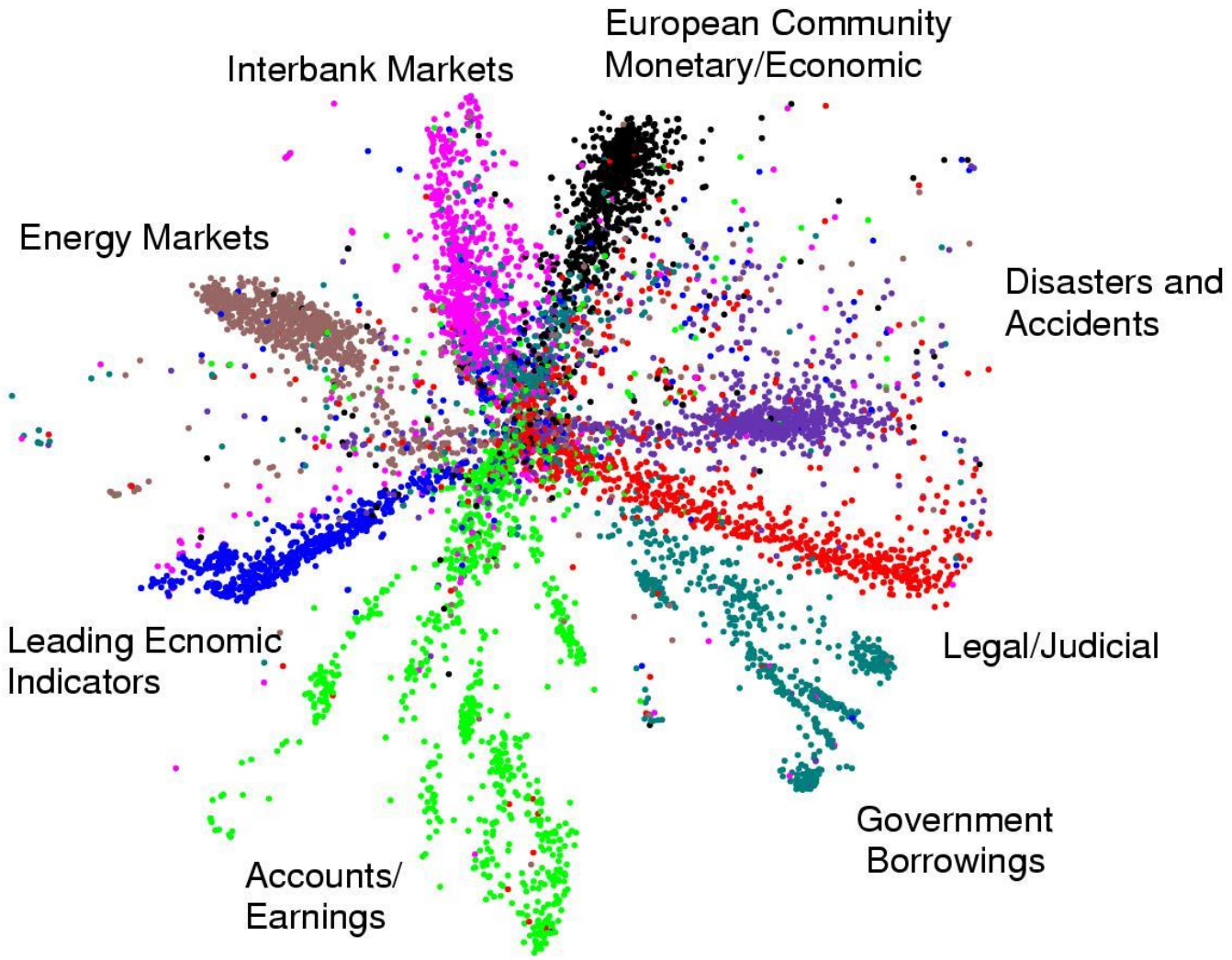
Supervised Learning: Uses

- **Prediction of future cases:** Use the rule to predict the output for future inputs
- **Knowledge extraction:** The rule is easy to understand
- **Compression:** The rule is simpler than the data it explains
- **Outlier detection:** Exceptions that are not covered by the rule, e.g., fraud

- **Unsupervised Learning:**
  - Unsupervised learning involves training the algorithm on an unlabeled dataset, and the algorithm must discover patterns or relationships in the data without explicit guidance on the correct output.
  - **Objective:** The goal is often to find hidden structures, group similar data points, or reduce the dimensionality of the data.
  - **Examples:** Clustering (e.g., grouping similar customers), dimensionality reduction (e.g., principal component analysis), and generative modeling (e.g., generating new data).

- **Clustering:**
  - Clustering involves grouping similar instances together based on some similarity or distance metric.
  - **Example:** K-means clustering for grouping customers based on purchasing behavior.

Displaying the structure of a set of documents

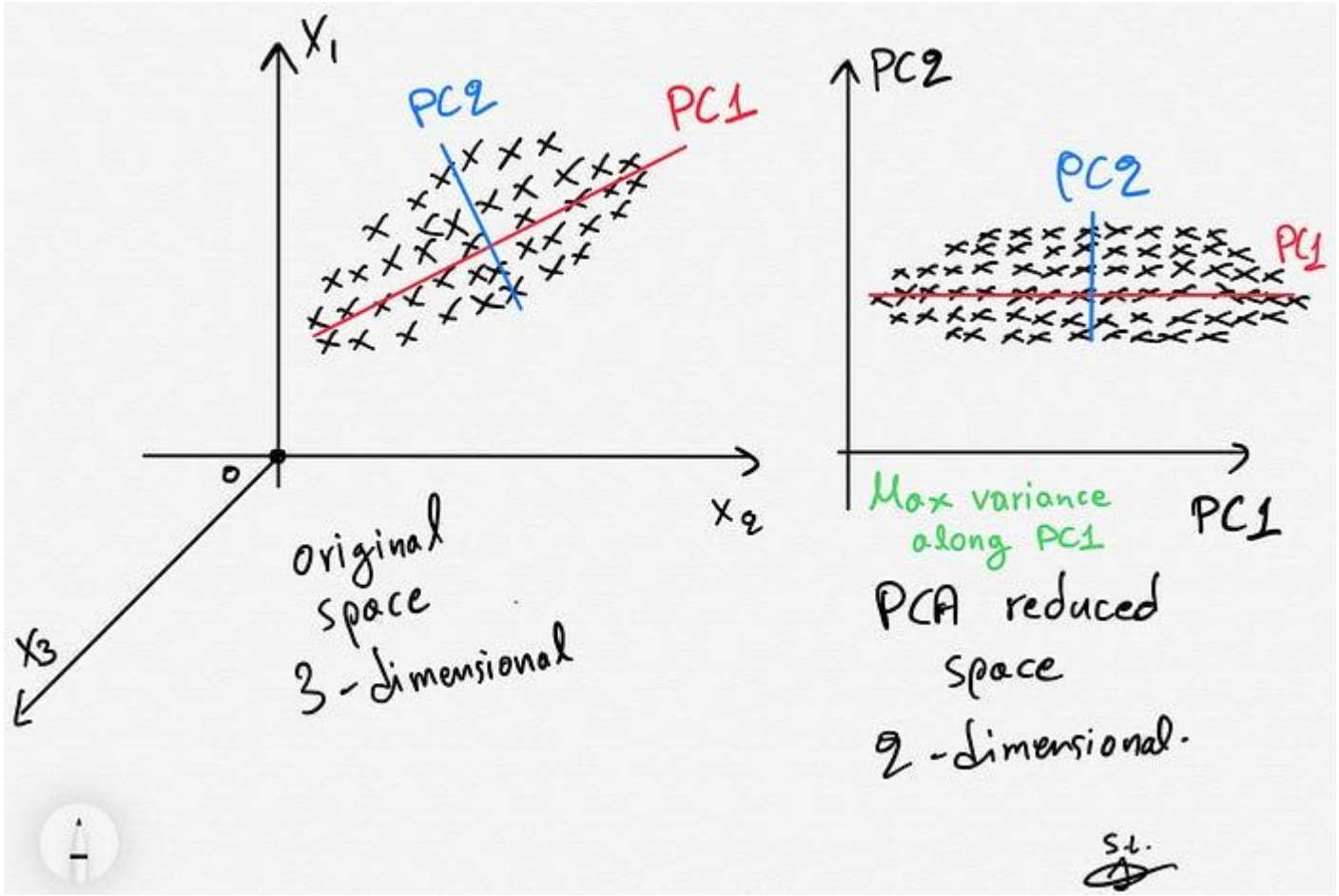


- **Dimensionality Reduction:**

  - Dimensionality reduction techniques aim to reduce the number of features in the data while preserving its essential characteristics.

  - **Example:** Principal Component Analysis (PCA) for reducing the dimensionality of high-dimensional data.

Dimensionality Reduction using PCA



- **Association:**
  - Association rule learning focuses on discovering interesting relationships or patterns in data, typically in the form of rules.
  - **Example:** Market basket analysis to identify associations between products frequently purchased together.

Learning Associations

- Basket analysis:

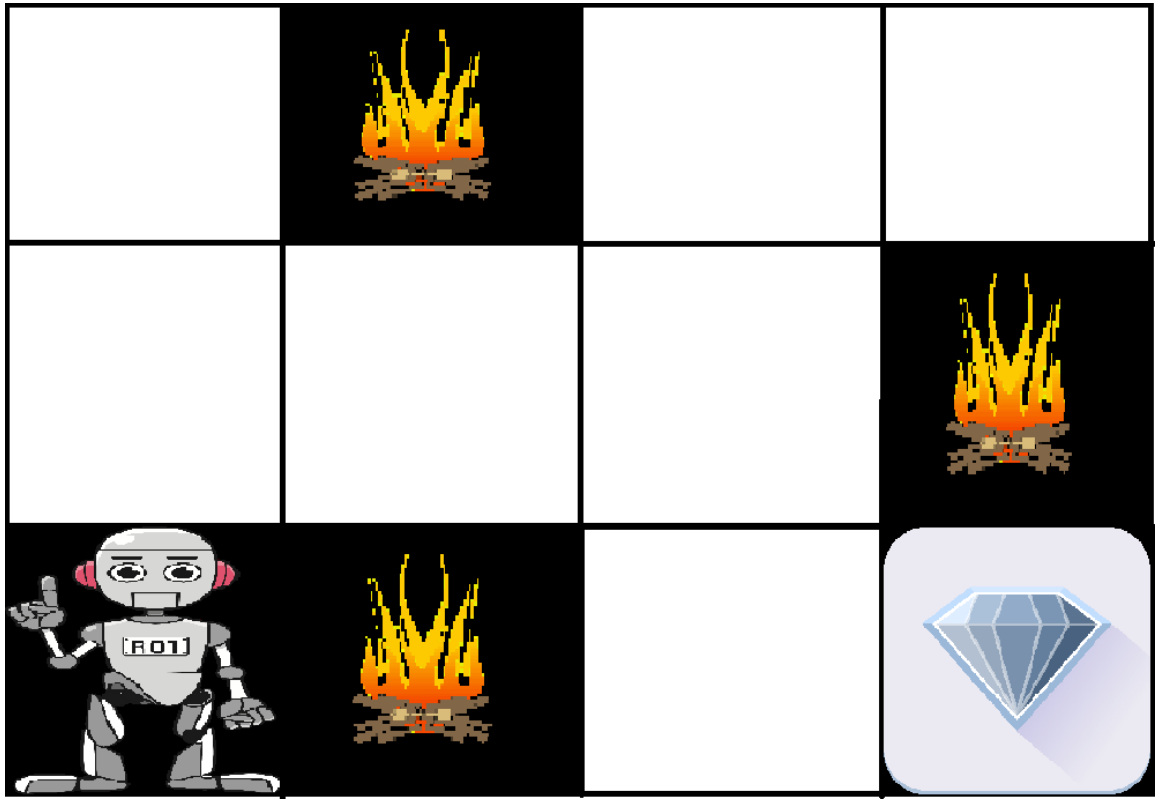
  $P(Y \mid X)$ probability that somebody who buys X also buys Y where X and Y are products/services.

  Example: $P(\text{chips} \mid \text{beer}) = 0.7$

- **Reinforcement Learning:**
  - Reinforcement learning involves an agent that learns to make decisions by interacting with an environment. The agent receives feedback in the form of rewards or penalties based on the actions it takes.
  - **Objective:** The goal is for the agent to learn a policy that maximizes cumulative reward over time.
  - **Examples:** Game playing (e.g., AlphaGo), robotic control, and autonomous systems.

Example

- We have an agent and a reward, with many hurdles in between. The agent is supposed to find the best possible path to reach the reward.
- The robot learns by trying all the possible paths and then choosing the path which gives him the reward with the least hurdles.



Other types of Learning

- **Semi-Supervised Learning:**
  - A combination of supervised and unsupervised learning where the algorithm is trained on a dataset that contains both labeled and unlabeled data.
- **Transfer Learning:**
  - The idea of training a model on one task and transferring its learned knowledge to another related task, potentially saving time and resources.
- **Ensemble Learning:**
  - A technique that involves combining multiple models (ensemble) to improve overall performance and robustness