

# Foundations of Machine Learning

## AI2000 and AI5000

FoML-01

What is ML? Types of Learning Paradigms

Dr. Konda Reddy Mopuri

Department of AI, IIT Hyderabad

July-Nov 2025



భారతీయ సాంకేతిక విజ్ఞాన సంస్థ హైదరాబాద్  
भारतीय प्रौद्योगिकी संस्थान हैदराबाद  
Indian Institute of Technology Hyderabad



# What is ML?

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 M. Mitchell, 1998

# Example Machine Learning problem

- Program for spam detection

# Example Machine Learning problem

- Program for spam detection
  - Task T: classifying an email into spam vs. non-spam



# Example Machine Learning problem

- Program for spam detection
  - Task T: classifying an email into spam vs. non-spam
  - Experience E: observing the users label the emails as spam vs. non-spam



# Example Machine Learning problem

- Program for spam detection
  - Task T: classifying an email into spam vs. non-spam
  - Experience E: observing the users label the emails as spam vs. non-spam
  - Performance P: fraction of email that are correctly detected

# Experience



భారతీయ సాంకేతిక విజ్ఞాన సంస్థ హైదరాబాద్  
भारतीय प्रौद्योगिकी संस्थान हैदराबाद  
Indian Institute of Technology Hyderabad



# Experience

Dear Beloved Friend,

I know this message will come to you as surprised but permit me of my desire to go into business relationship with you.

I am Miss [REDACTED] a daughter to late [REDACTED] of Libya whom was murdered during the recent civil war in Libya in March 2011, before his death my late father was a strong supporter and a member of late Moammar Gadhafi Government in Tripoli. Meanwhile before the incident, my late Father came to Cotonou Benin republic with the sum of USD4, 200,000.00 (US\$4.2M) which he deposited in a Bank here in Cotonou Benin Republic West Africa for safe keeping.

I am here seeking for an avenue to transfer the fund to you in only you're reliable and trustworthy person to Investment the fund. I an here in Benin Republic because of the death of my parent's and I want you to help me transfer the fund into your bank account for investment purpose.

Please I will offer you 20% of the total sum of USD4.2M for your assistance. Please I wish to transfer the fund urgently without delay into your account and also wish to relocate to your country due to the poor condition in Benin, as to enable me continue my education as I was a medical student before the sudden death of my parent's. Reply to my alternative email: [REDACTED], Your immediate response would be appreciated.

Remain blessed,

Miss [REDACTED]





# Experience

Dear Beloved Friend,

I know this message will come to you as surprised but permit me of my desire to go into business relationship with you.

I am Miss [REDACTED] a daughter to late [REDACTED] of Libya whom was murdered during the recent civil war in Libya in March 2011, before his death my late father was a strong supporter and a member of late Moammar Gadhafi Government in Tripoli. Meanwhile before the incident, my late Father came to Cotonou Benin republic with the sum of USD4, 200,000.00 (US\$4.2M) which he deposited in a Bank here in Cotonou Benin Republic West Africa for safe keeping.

I am here seeking for an avenue to transfer the fund to you in only you're reliable and trustworthy person to Investment the fund. I an here in Benin Republic because of the death of my parent's and I want you to help me transfer the fund into your bank account for investment purpose.

Please I will offer you 20% of the total sum of USD4.2M for your assistance. Please I wish to transfer the fund urgently without delay into your account and also wish to relocate to your country due to the poor condition in Benin, as to enable me continue my education as I was a medical student before the sudden death of my parent's. Reply to my alternative email: [REDACTED], Your immediate response would be appreciated.

Remain blessed,

Miss [REDACTED]



# Experience is through the data



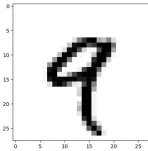
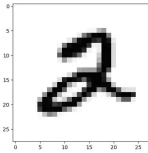
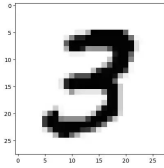
# T: Class of tasks

- Classification



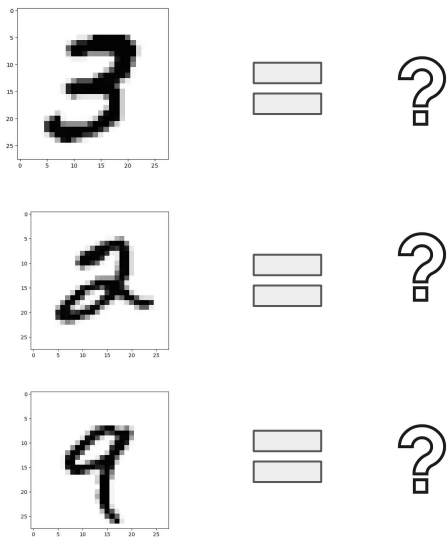
# T: Class of tasks

- Classification



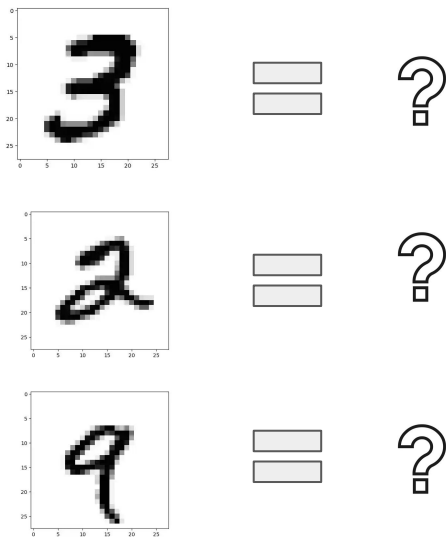
# T: Class of tasks

- Classification



# T: Class of tasks

- Classification



Discrete outputs/targets  
 $\{0, 1, 2, 3, \dots, 9\}$



# T: Class of tasks

- Regression



# T: Class of tasks

- Regression

Input  $x$





# T: Class of tasks

- Regression

Input  $x$

Target  $t = \sin(2\pi x) + \epsilon$

Noise  $\epsilon \sim N(0, 1)$



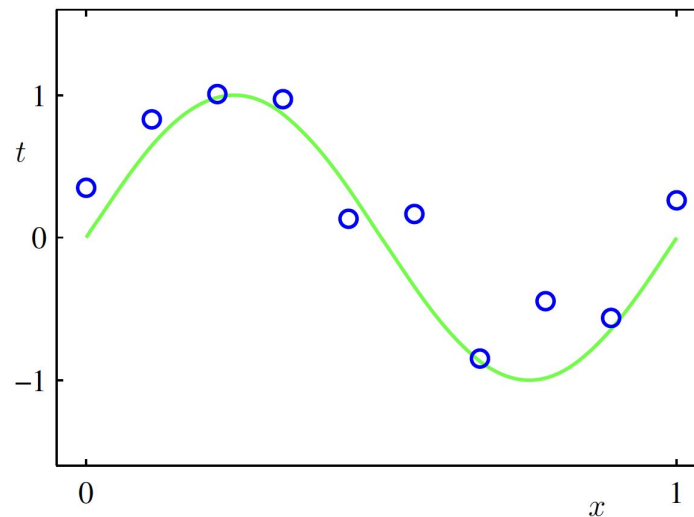
# T: Class of tasks

- Regression

Input  $x$

Target  $t = \sin(2\pi x) + \epsilon$

Noise  $\epsilon \sim N(0, 1)$



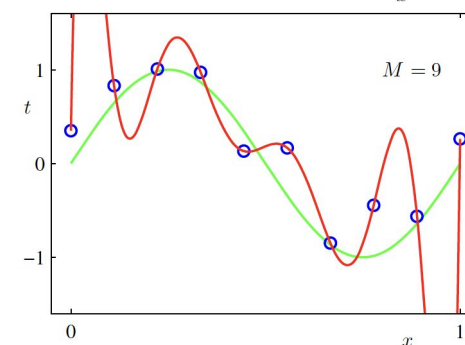
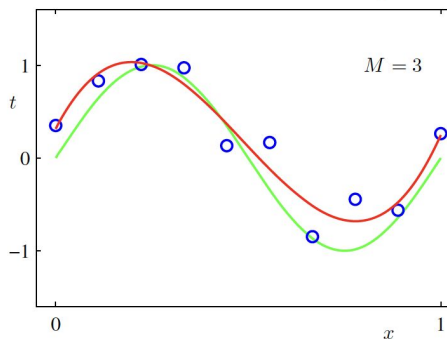
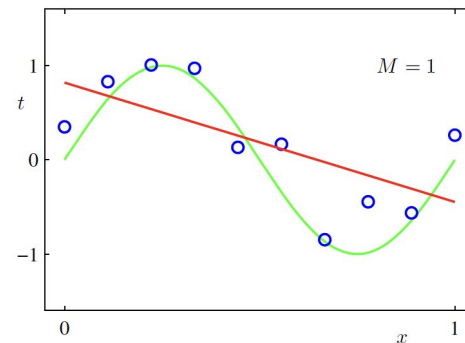
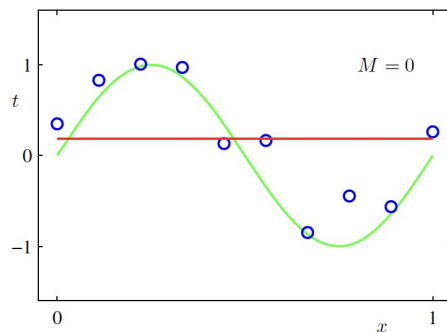
# T: Class of tasks

- Regression

Input  $x$

Target  $t = \sin(2\pi x) + \epsilon$

Noise  $\epsilon \sim N(0, 1)$



# T: Class of tasks

- Clustering

# T: Class of tasks

- Clustering

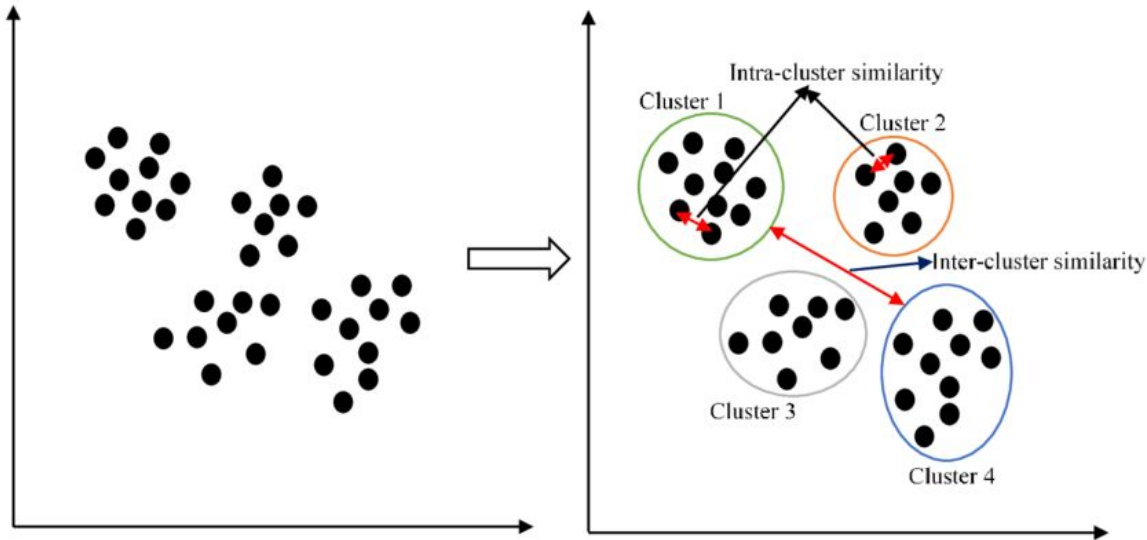


Figure credits - [Absalom El-Shamir Ezugwu](#)

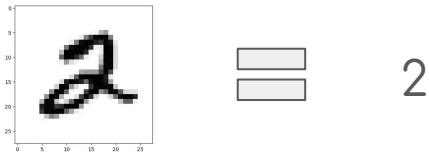
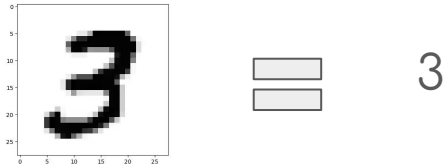
# P: Performance Measure

- Classification



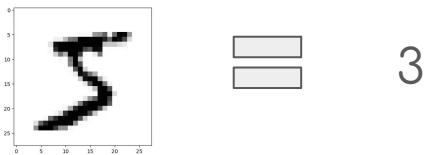
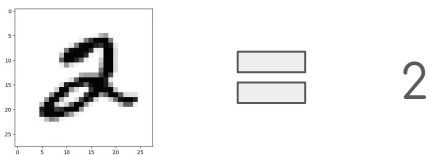
# P: Performance Measure

- Classification



# P: Performance Measure

- Classification



$$\text{Accuracy}(\mathbf{t}, \hat{\mathbf{t}}) = \frac{1}{N} \sum_{i=1}^N \mathbb{1}(\mathbf{t}_i = \hat{\mathbf{t}}_i)$$





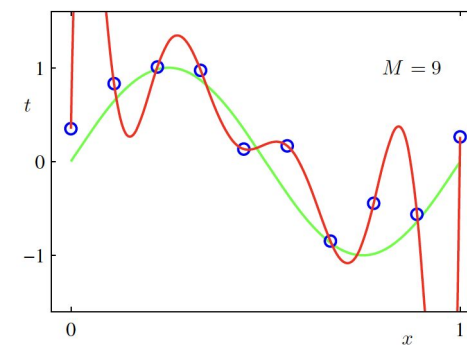
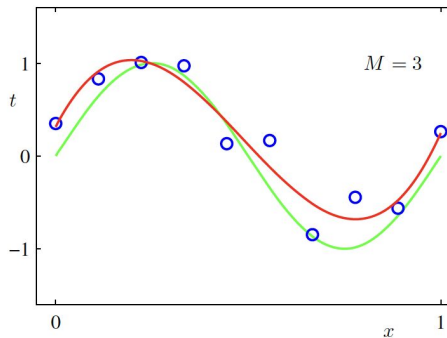
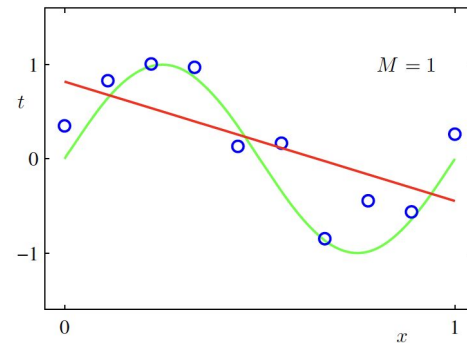
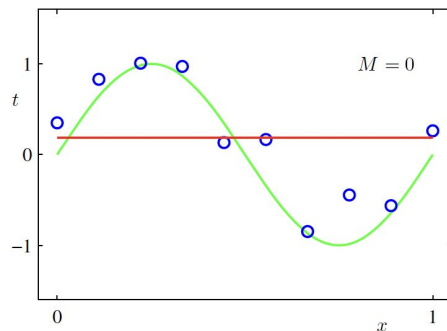
# P: Performance Measure

- Regression



# P: Performance Measure

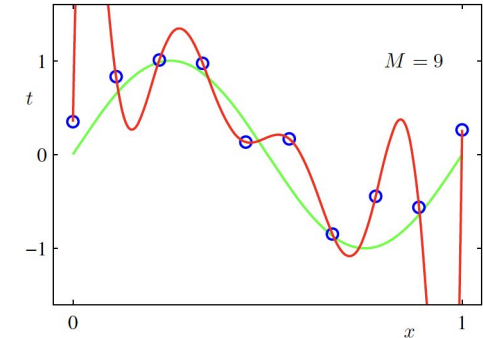
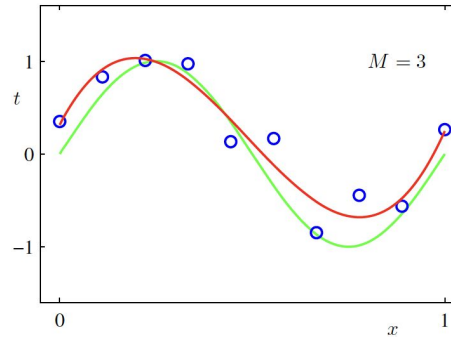
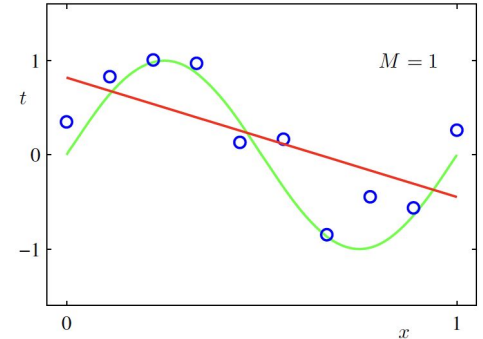
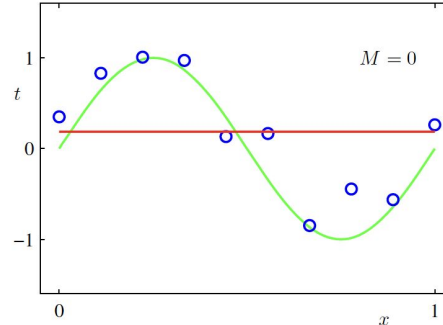
- Regression



# P: Performance Measure

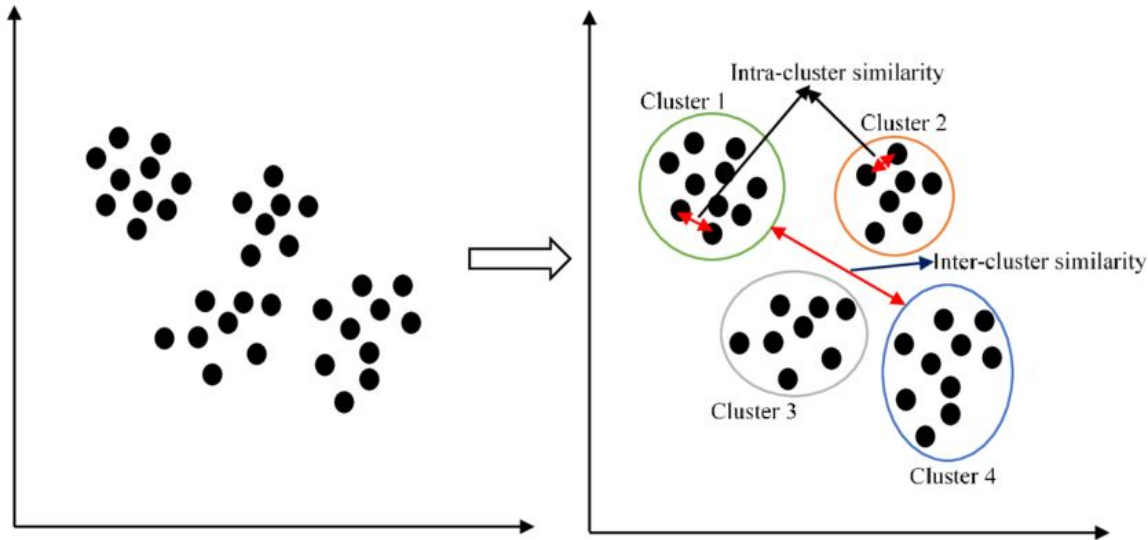
- Regression

$$\text{MSE}(\mathbf{t}, \hat{\mathbf{t}}) = \frac{1}{N} \sum_{i=1}^N \mathbb{1}(\mathbf{t}_i - \hat{\mathbf{t}}_i)^2$$



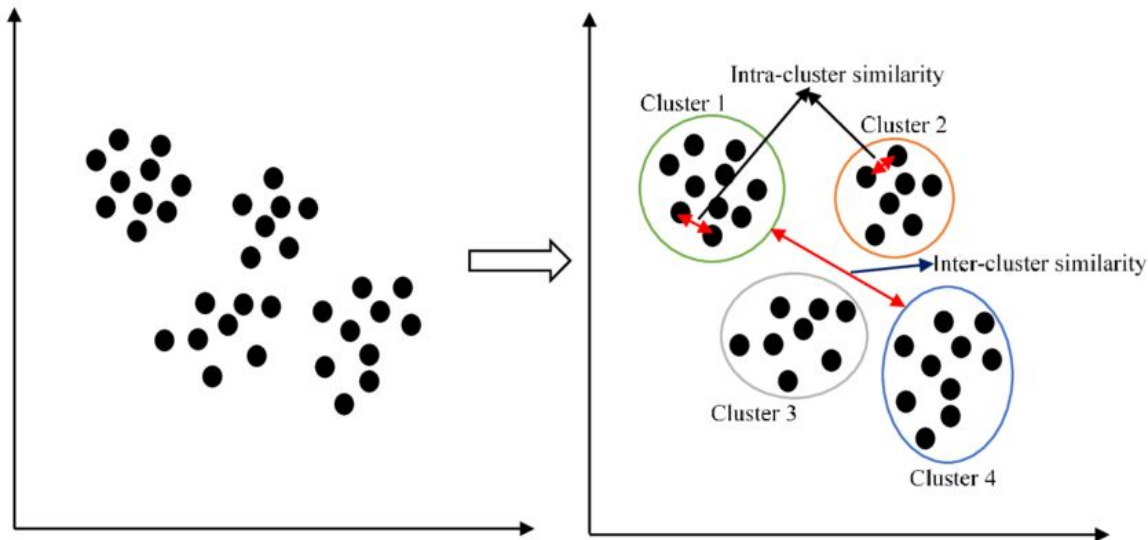
# P: Performance Measure

- Clustering



# P: Performance Measure

- Clustering



$$WCSS = \sum_{i=1}^N \min_{\mu_j \in C} (\mu_j - x_i)^2$$



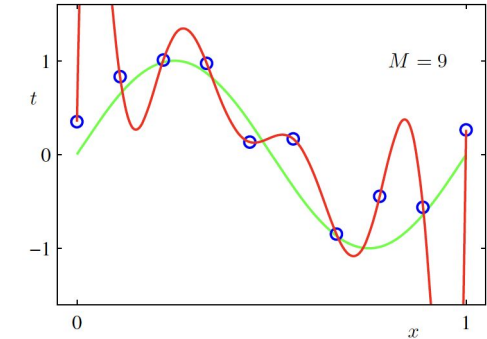
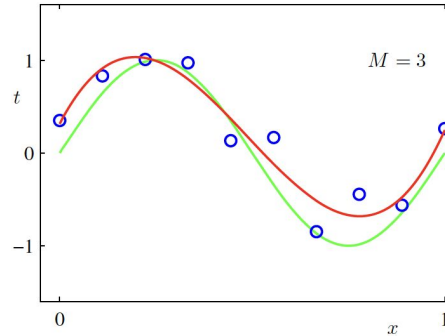
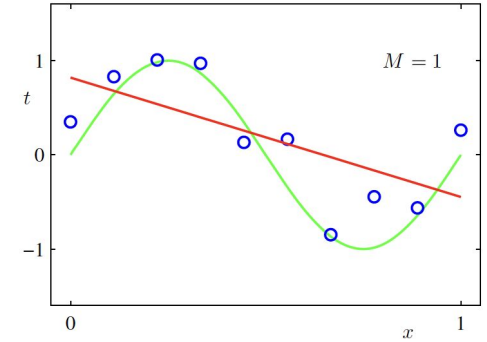
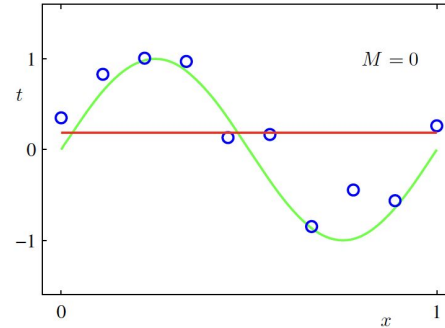
# P: Performance Measure

- On which samples it should be measured?



# P: Performance Measure

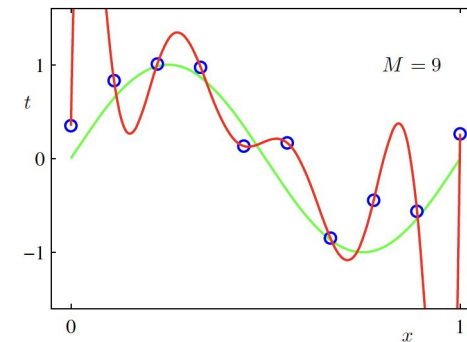
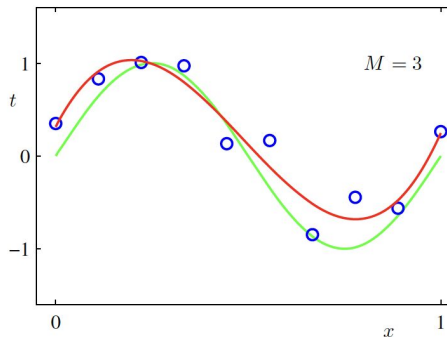
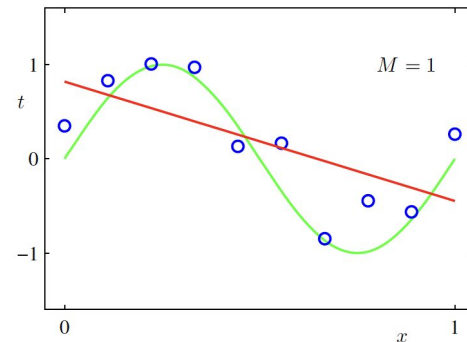
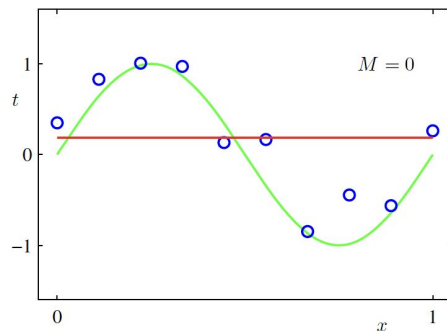
- On which samples it should be measured?



# P: Performance Measure

- On which samples it should be measured?

$$\text{MSE}(\mathbf{t}, \hat{\mathbf{t}}) = \frac{1}{N} \sum_{i=1}^N \mathbb{1}(\mathbf{t}_i - \hat{\mathbf{t}}_i)^2$$

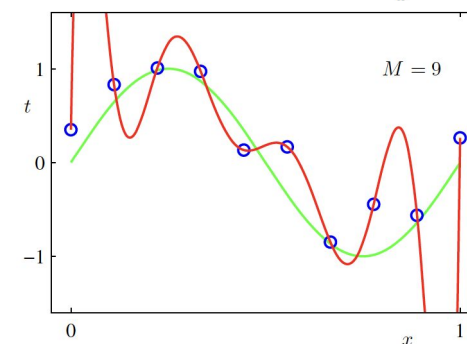
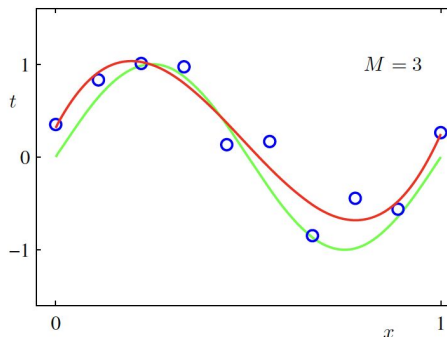
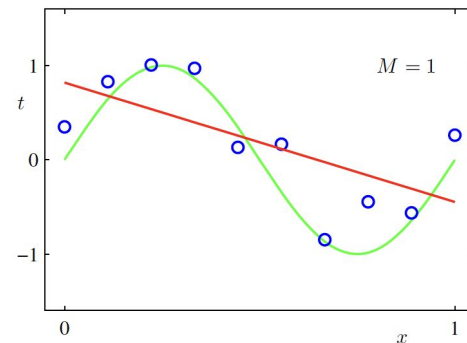
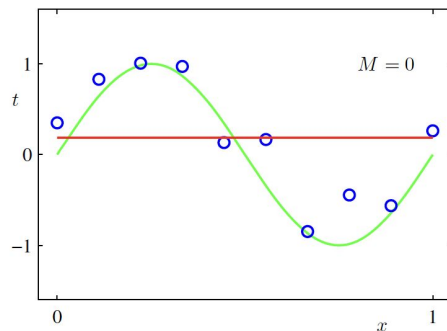




# P: Performance Measure

- Best performance on train set?

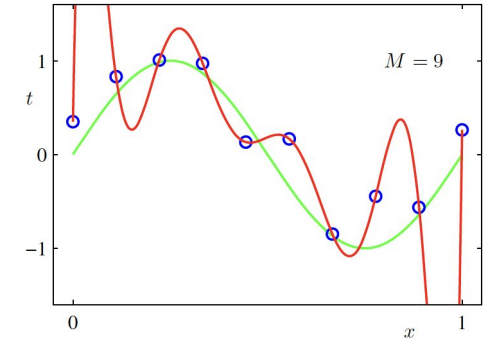
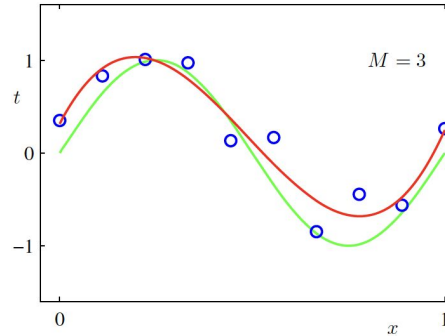
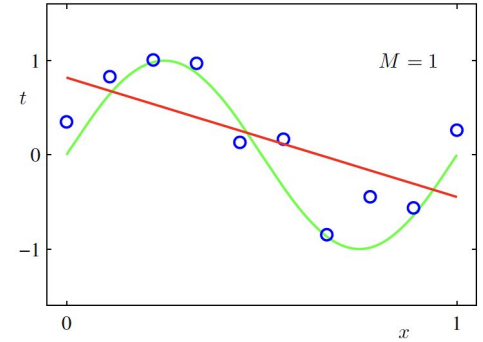
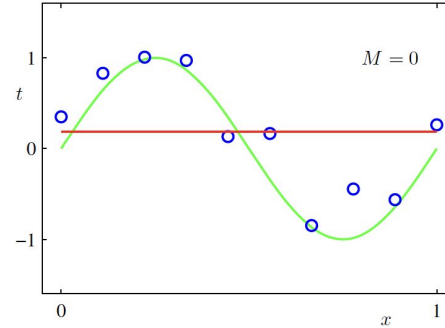
$$\text{MSE}(\mathbf{t}, \hat{\mathbf{t}}) = \frac{1}{N} \sum_{i=1}^N \mathbb{1}(\mathbf{t}_i - \hat{\mathbf{t}}_i)^2$$



# P: Performance Measure

- Best performance on new samples?

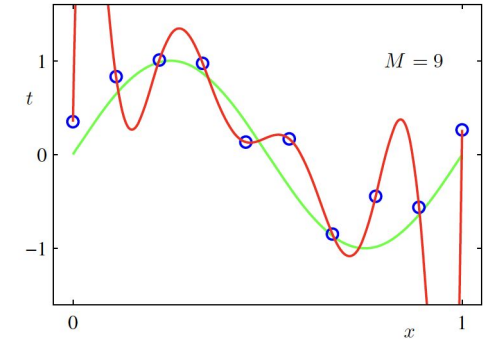
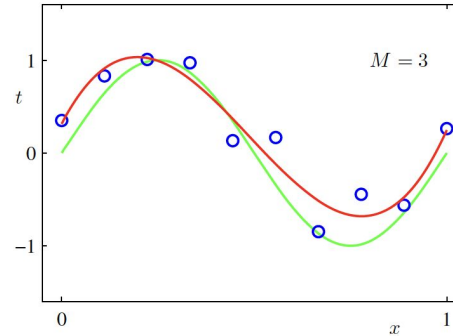
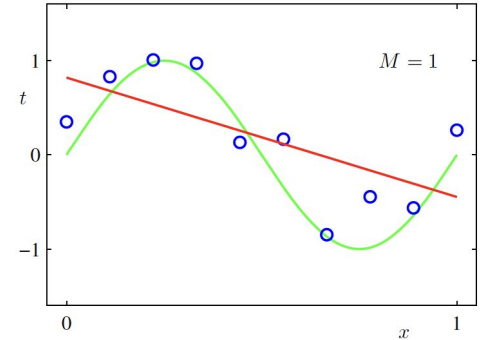
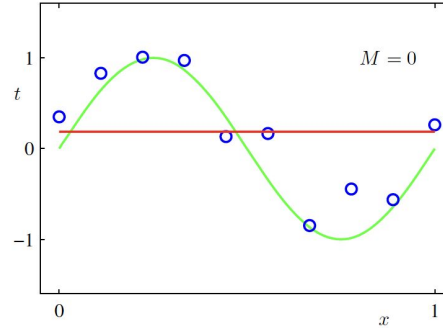
$$\text{MSE}(\mathbf{t}, \hat{\mathbf{t}}) = \frac{1}{N} \sum_{i=1}^N \mathbb{1}(\mathbf{t}_i - \hat{\mathbf{t}}_i)^2$$



# P: Performance Measure

- Best performance on new samples?

$$\text{MSE}(\mathbf{t}, \hat{\mathbf{t}}) = \frac{1}{N} \sum_{i=1}^N \mathbb{1}(\mathbf{t}_i - \hat{\mathbf{t}}_i)^2$$

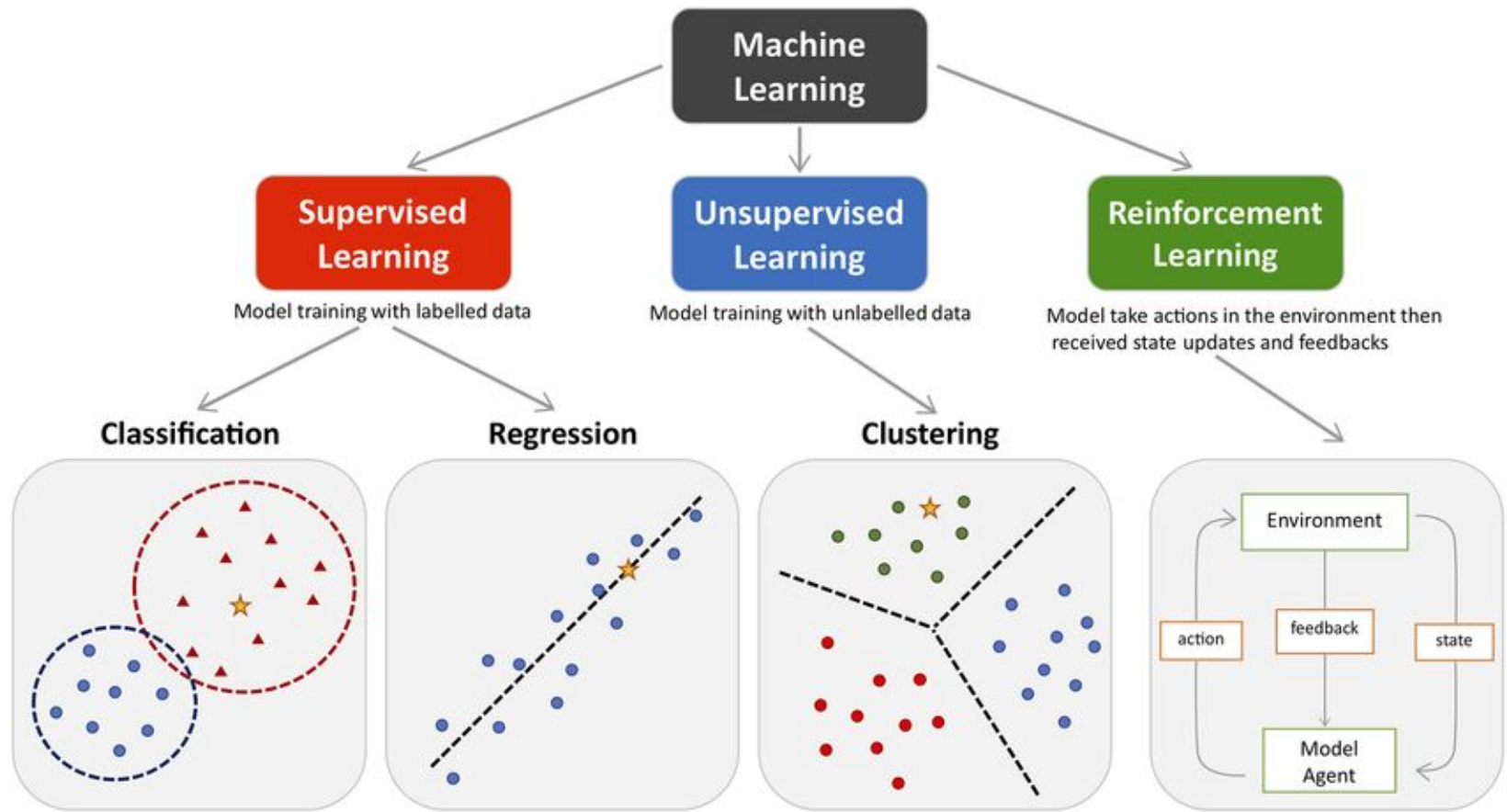


Generalization



# Types of Machine Learning





# Supervised Learning



భారతీయ సాంకేతిక విజ్ఞాన సంస్థ హైదరాబాద్  
भारतीय प्रौद्योगिकी संस्थान हैदराबाद  
Indian Institute of Technology Hyderabad



# Supervised Learning

- Supervised learning techniques
  - $p$  features  $X^1, X^2, X^3, \dots, X^p$  measured on  $N$  observations

# Supervised Learning

- Supervised learning techniques
  - $p$  features  $X^1, X^2, X^3, \dots, X^p$  measured on  $N$  observations
  - Response ( $t$ ) also measured on these

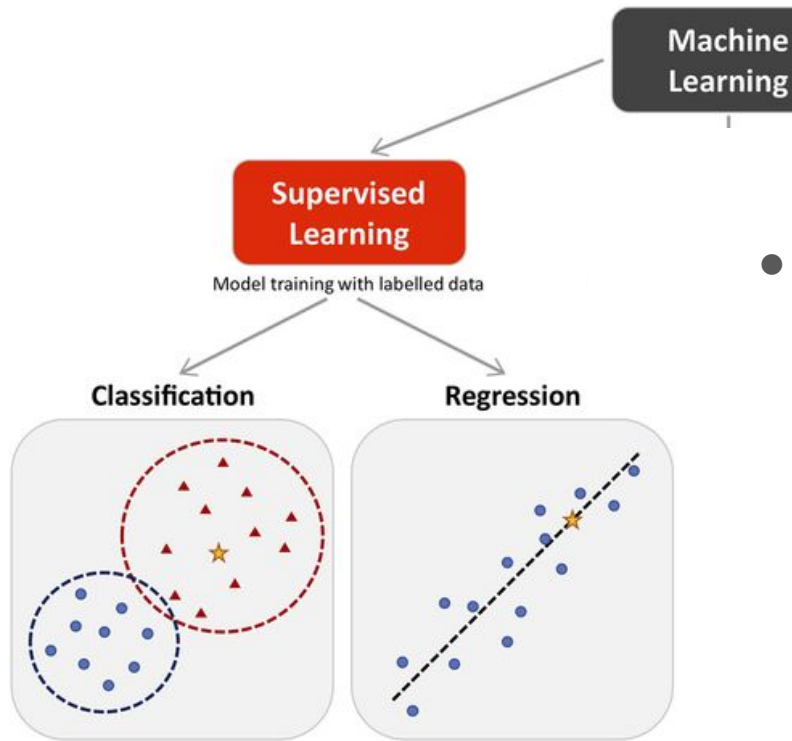


# Supervised Learning

- Supervised learning techniques
  - $p$  features  $X^1, X^2, X^3, \dots, X^p$  measured on  $N$  observations
  - Response ( $t$ ) also measured on these
  - $\rightarrow$  goal is to predict  $t$  using  $X^1, X^2, X^3, \dots, X^p$



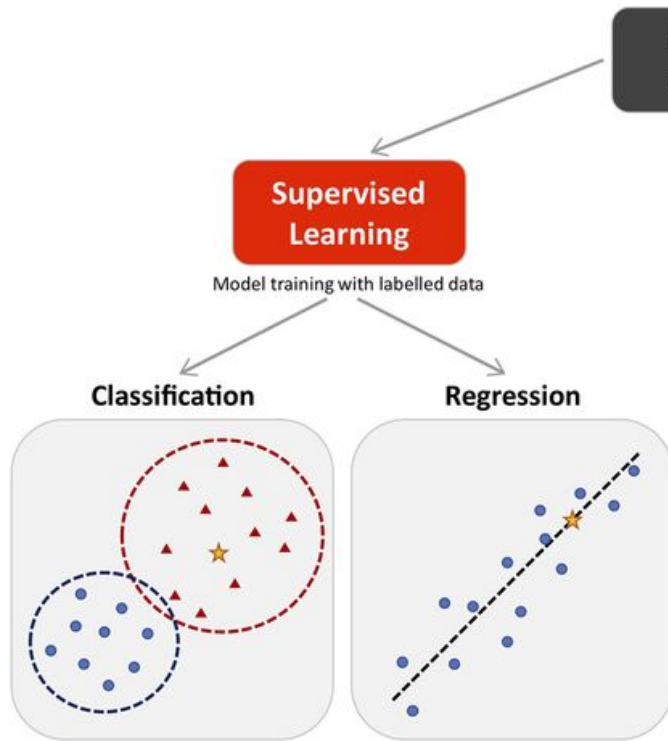
# Supervised Learning



- Dataset: {features  $x_i$ , targets  $t_i$ }



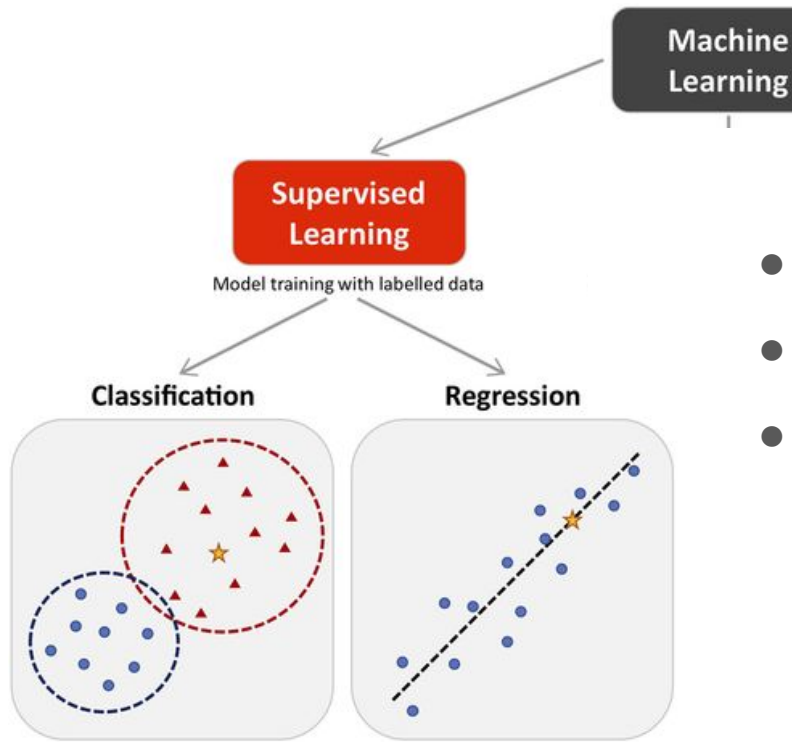
# Supervised Learning



- Dataset: {features  $x_i$ , targets  $t_i$ }
- Classification: discrete targets



# Supervised Learning



- Dataset: {features  $x_i$ , targets  $t_i$ }
- Classification: discrete targets
- Regression: continuous targets



# Supervised Learning

Finding a function  $f$  such that  $f(x) \approx t$  for all known and unknown samples  $(x, t)$



# Unsupervised learning



భారతీయ సాంకేతిక విజ్ఞాన సంస్థ హైదరాబాద్  
भारतीय प्रौद्योगिकी संस्थान हैदराबाद  
Indian Institute of Technology Hyderabad



# Unsupervised learning

- Only have a set of features  $X^1, X^2, X^3, \dots, X^p$



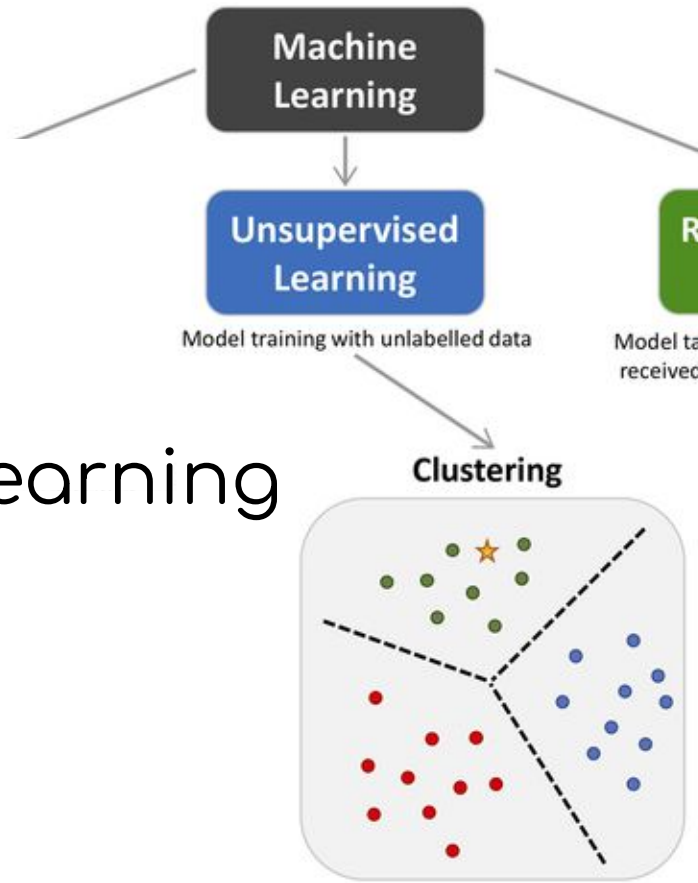
# Unsupervised learning

- Only have a set of features  $X^1, X^2, X^3, \dots, X^p$
- Not interested in prediction (don't have an associated  $t$ )



# Unsupervised learning

- Only have a set of features  $X^1, X^2, X^3, \dots, X^p$
- Not interested in prediction (don't have an associated  $t$ )
- $\rightarrow$  goal is to discover “Interesting things” about the data



# Unsupervised Learning

# Unsupervised learning

- “Interesting things” about the data



# Unsupervised learning

- “Interesting things” about the data
  - Is there an informative way to visualize the data?



# Unsupervised learning

- “Interesting things” about the data
  - Is there an informative way to visualize the data?
  - Can we discover ‘subgroups’ among the variables or samples?

# Unsupervised learning

- “Interesting things” about the data
  - Is there an informative way to visualize the data?
  - Can we discover ‘subgroups’ among the variables or samples?
  - Can we compress the data?

# Unsupervised learning

- A diverse set of statistical techniques for answering such questions
  - Clustering



# Unsupervised learning

- A diverse set of statistical techniques for answering such questions
  - Clustering
  - Dimensionality Reduction - Principal Component Analysis (PCA)



# Reinforcement Learning

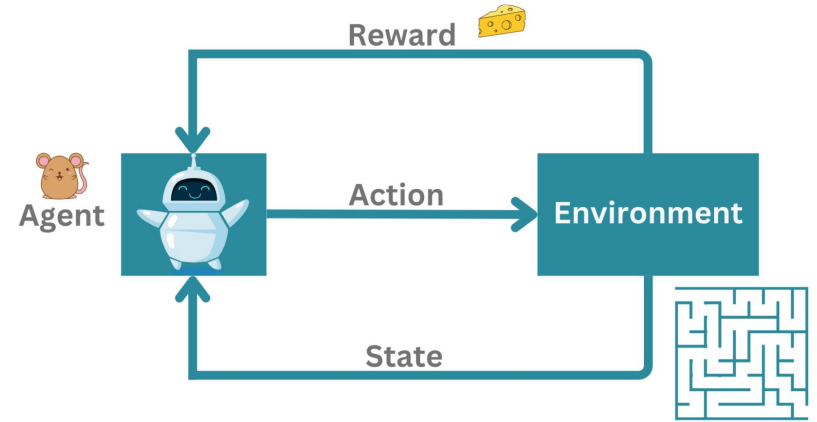


భారతీయ సాంకేతిక విజ్ఞాన సంస్థ హైదరాబాద్  
भारतीय प्रौद्योगिकी संस्थान हैदराबाद  
Indian Institute of Technology Hyderabad



# Reinforcement Learning

- Dynamic environment → provides its state information

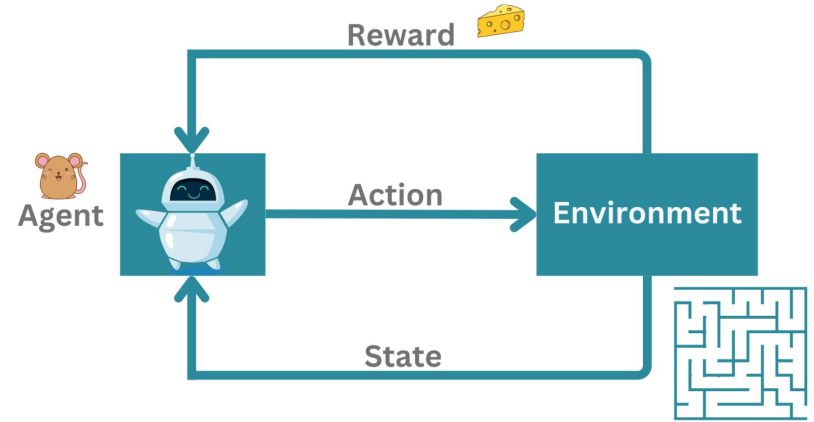


<https://www.kdnuggets.com/>



# Reinforcement Learning

- Dynamic environment → provides its state information
- Agent → takes actions and receives rewards

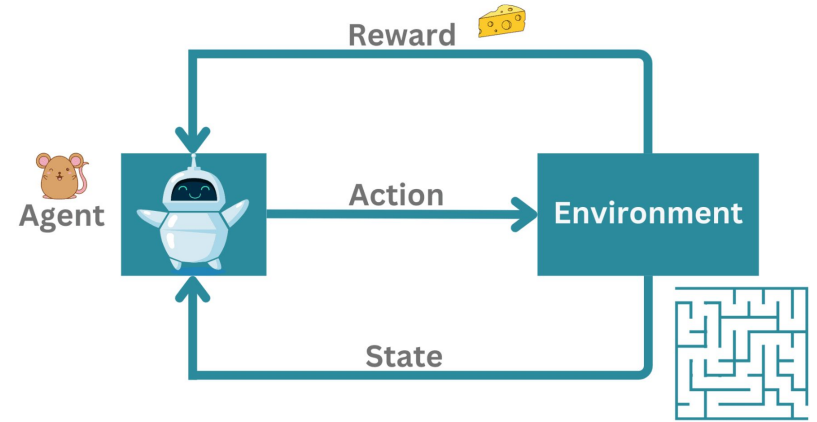


<https://www.kdnuggets.com/>



# Reinforcement Learning

- Task: maximizing the reward

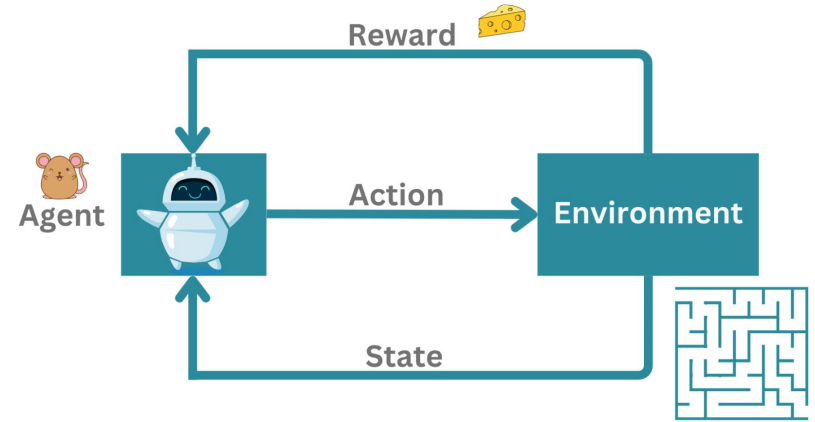


<https://www.kdnuggets.com/>



# Reinforcement Learning

- Task: maximizing the reward
- Learning by trial-and-error

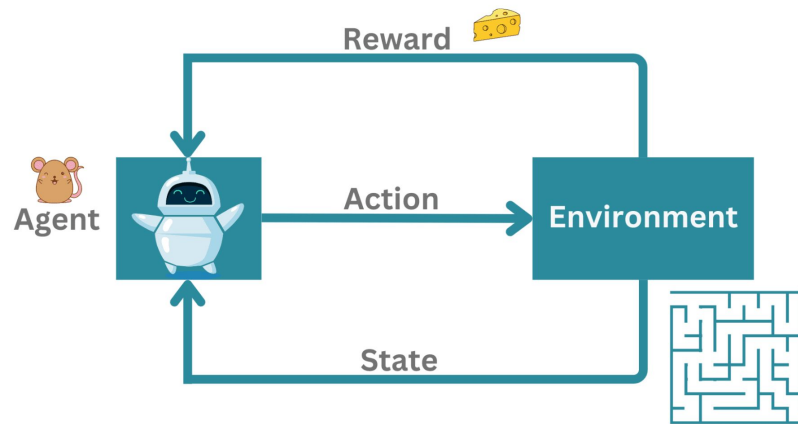


<https://www.kdnuggets.com/>



# Reinforcement Learning

- Task: maximizing the reward
- Learning by trail-and-error
- Games, Robotics, etc.



<https://www.kdnuggets.com/>



# Other types of learning

- Semi-supervised learning



# Other types of learning

- Semi-supervised learning
  - Data -  $\{X_1, X_2, \dots, X_N\}$





# Other types of learning

- Semi-supervised learning
  - Data -  $\{X_1, X_2, \dots, X_N\}$
  - Target -  $\{t_1, t_2, \dots, t_K\}$



# Other types of learning

- Semi-supervised learning
  - Data -  $\{X_1, X_2, \dots, X_N\}$
  - Target -  $\{t_1, t_2, \dots, t_K\}$
  - $K < N$  (Not all samples have the labels!)



# Other types of learning

- Semi-supervised learning
  - Data -  $\{X_1, X_2, \dots, X_N\}$
  - Target -  $\{t_1, t_2, \dots, t_K\}$
  - $K < N$  (Not all samples have the labels!)
  - Work with all the data to learn the predictor



# Next

## Probability - Bayes Theorem, Expectation, Variance

