



उत्तिष्ठत आगत्य पुनश्च वरदानं विनोदयत
आयुः शान्तिः सौख्यं धनं सौभाग्यं सन्तानं
आरोग्यं सौमित्राणां प्रियं सौ शत्रूणां शत्रुः, विष्णुः

1/ 44

Vishal Kumar
Jaiswal

Introduction

Machine Learning

Fundamentals

Inductive learning

Supervised Learning

Unsupervised
Learning

Reinforcement
Learning

Survey
Method

Mathematical
Preliminaries

Learning
Algorithms

Entropy

Ensemble
Learning
Method

Machine Learning using Python

Theory and Implementation

Vishal Kumar Jaiswal

M.Tech

CST Department , IIST Shibpur

June 18, 2016

Summer Internship 2016



Table of Contents

2/ 44

Vishal Kumar
Jaiswal

Introduction

Machine Learning

Fundamentals

Inductive learning

Supervised Learning

Unsupervised
Learning

Reinforcement
Learning

Survey
Method

Mathematical
Preliminaries

Learning
Algorithms

Entropy

Ensemble
Learning
Method

- 1 Introduction
 - Machine Learning
- 2 Fundamentals
 - Inductive learning
 - Supervised Learning
 - Unsupervised Learning
 - Reinforcement Learning
- 3 Survey Method
- 4 Mathematical Preliminaries
- 5 Learning Algorithms
- 6 Entropy
- 7 Ensemble Learning Method



What to expect

3/ 44

Vishal Kumar
Jaiswal

Introduction

Machine Learning

Fundamentals

Inductive learning

Supervised Learning

Unsupervised
Learning

Reinforcement
Learning

Survey Method

Mathematical Preliminaries

Learning Algorithms

Entropy

Ensemble Learning Method

- You'll learn by doing,
- Bring machine learning to life using specified use case,
- Identify financial fraud by mining email datasets to identify writer of email.

Why machine learning?

4/ 44

Vishal Kumar
Jaiswal

Introduction

Machine Learning

Fundamentals

Inductive learning

Supervised Learning

Unsupervised
Learning

Reinforcement
Learning

Survey
Method

Mathematical
Preliminaries

Learning
Algorithms

Entropy

Ensemble
Learning
Method

- Automating automation
- Getting computers to program themselves
- Writing software is the bottleneck
- Let the data do the work instead!
- Can help in solve the most practical society impacting problems.

Comparison with traditional programming

5/ 44

Vishal Kumar
Jaiswal

Introduction

Machine Learning

Fundamentals

Inductive learning

Supervised Learning

Unsupervised

Learning

Reinforcement

Learning

Survey
Method

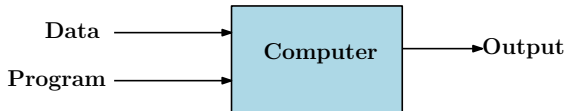
Mathematical
Preliminaries

Learning
Algorithms

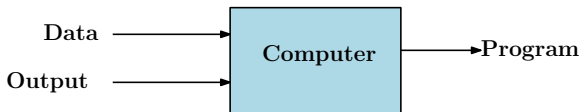
Entropy

Ensemble
Learning
Method

Traditional Programming:-



Machine Learning:-





Machine learning

6/ 44

Vishal Kumar
Jaiswal

Introduction

Machine Learning

Fundamentals

Inductive learning

Supervised Learning

Unsupervised
Learning

Reinforcement
Learning

Survey
Method

Mathematical
Preliminaries

Learning
Algorithms

Entropy

Ensemble
Learning
Method

- Building computational artifacts that learn over time based on experience.
- Employed in computing tasks where designing and programming explicit algorithms is infeasible.
- Build a model from example inputs in order to make data-driven predictions or decisions expressed as outputs rather than following strictly static program instructions.

Process of **learning by example** - where a system tries to induce a general rule from a set of observed instances i.e. specific to general rule.

- 1 **Discrete:-** Classification
- 2 **Continuous:-** Regression

Constructing class definitions is called **Inductive learning**.

Note:

Artificial Intelligence is based on deductive learning.

Process of **learning by example** - where a system tries to induce a general rule from a set of observed instances i.e. specific to general rule.

- ① **Discrete:-** Classification
- ② **Continuous:-** Regression

Constructing class definitions is called **Inductive learning**.

Note:

Artificial Intelligence is based on deductive learning.

Classification of Machine Learning

8/ 44

Vishal Kumar
Jaiswal

Introduction

Machine Learning

Fundamentals

Inductive learning

Supervised Learning

Unsupervised
Learning

Reinforcement
Learning

Survey
Method

Mathematical
Preliminaries

Learning
Algorithms

Entropy

Ensemble
Learning
Method

Based on feedback available to learning system, we categorize machine learning algorithms in following categories:-

- Supervised learning
- Unsupervised learning
- Reinforcement learning



Machine Learning: Type 1

9/ 44

Vishal Kumar
Jaiswal

Introduction

Machine Learning

Fundamentals

Inductive learning

Supervised Learning

Unsupervised
Learning

Reinforcement
Learning

Survey
Method

Mathematical
Preliminaries

Learning
Algorithms

Entropy

Ensemble
Learning
Method

Supervised learning

Function approximation from input and output pairs (training dataset) i.e. generalization.

Supervised Learning

10/ 44

Vishal Kumar
Jaiswal

Introduction

Machine Learning

Fundamentals

Inductive learning

Supervised Learning

Unsupervised
Learning

Reinforcement
Learning

Survey
Method

Mathematical
Preliminaries

Learning
Algorithms

Entropy

Ensemble
Learning
Method

Classification

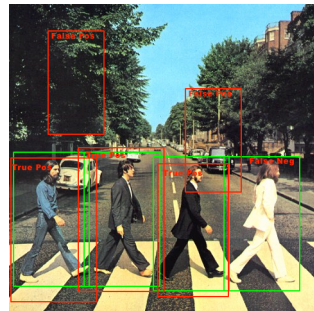
Assign, to a particular input, the name of class to which it belongs.

Classification: An Example

11/ 44

Vishal Kumar
Jaiswal

Introduction
 Machine Learning
 Fundamentals
 Inductive learning
 Supervised Learning
 Unsupervised Learning
 Reinforcement Learning
 Survey Method
 Mathematical Preliminaries
 Learning Algorithms
 Entropy
 Ensemble Learning Method



Spam filtering and Object detection

Supervised learning

12/ 44

Vishal Kumar
Jaiswal

Introduction

Machine Learning

Fundamentals

Inductive learning

Supervised Learning

Unsupervised
Learning

Reinforcement
Learning

Survey
Method

Mathematical
Preliminaries

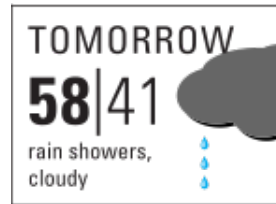
Learning
Algorithms

Entropy

Ensemble
Learning
Method

Regression

Predicting a numerical value for input dataset.



stock market prediction and weather forecasting



Machine learning: Type 2

14/ 44

Vishal Kumar
Jaiswal

Introduction

Machine Learning

Fundamentals

Inductive learning

Supervised Learning

Unsupervised
Learning

Reinforcement
Learning

Survey
Method

Mathematical
Preliminaries

Learning
Algorithms

Entropy

Ensemble
Learning
Method

Unsupervised learning

Without any examples, summarize or describe data just by looking at the input.

e.g. naming all animals dog by a child. Describe people based on ethnicity like Punjabi, Bangali, South Indian etc.



Machine learning: Type 2

14/ 44

Vishal Kumar
Jaiswal

Introduction

Machine Learning

Fundamentals

Inductive learning

Supervised Learning

Unsupervised
Learning

Reinforcement
Learning

Survey
Method

Mathematical
Preliminaries

Learning
Algorithms

Entropy

Ensemble
Learning
Method

Unsupervised learning

Without any examples, summarize or describe data just by looking at the input.

e.g. naming all animals dog by a child. Describe people based on ethnicity like Punjabi, Bangali, South Indian etc.

Unsupervised Learning

15/ 44

Vishal Kumar
Jaiswal

Introduction

Machine Learning

Fundamentals

Inductive learning

Supervised Learning

Unsupervised
Learning

Reinforcement
Learning

Survey
Method

Mathematical
Preliminaries

Learning
Algorithms

Entropy

Ensemble
Learning
Method

Clustering

Discovering structure in data e.g. finding similar images in google images, Collaborative filtering of news feed content by facebook.

Clustering: An Example

16/ 44

Vishal Kumar
Jaiswal

Introduction

Machine Learning

Fundamentals

Inductive learning

Supervised Learning

Unsupervised
Learning

Reinforcement
Learning

Survey
Method

Mathematical
Preliminaries

Learning
Algorithms

Entropy

Ensemble
Learning
Method

News

India edition ▾

Modern ▾

हिन्दी

தமிழ்

සමතුලිත

தெலுగ్

Personalise

Top Stories

Real Madrid C.F.
Zika virus
Narendra Modi
Sunrisers Hyderabad
NASA
Taliban
Brazil
Syria
Arun Jaitley
Virat Kohli

Howrah, West Bengal

Suggested for you

India

World

Top Stories



See realtime
coverage

Change has come, says PM Narendra Modi; vows to root out corruption

Economic Times - 7 hours ago



NEW DELHI: Celebrating second anniversary of his government, Prime Minister Narendra Modi today said a "change" has come in governance and vowed to root out corruption and make life easier for the people who have ...

Aam aadmi finds no place in Modi govt. celebrations The Hindu

A big bang theory The Indian Express



India Today

NDTV

India Today



Livemint

Modi govt celebrates 2 yrs in office with a gala event

India Today - 8 hours ago

New Delhi, May 28 (PTI) The Narendra Modi government today celebrated its two years in office with a gala event, which had a smattering megastar Amitabh Bachchan where the Prime Minister vowed to root out corruption and ...

Google news

Machine learning: Type 3

17/ 44

Vishal Kumar
Jaiswal

Introduction

Machine Learning

Fundamentals

Inductive learning

Supervised Learning

Unsupervised
Learning

Reinforcement
Learning

Survey
Method

Mathematical
Preliminaries

Learning
Algorithms

Entropy

Ensemble
Learning
Method

Reinforcement learning

Learn how to behave based on delayed feedback from the environment.

e.g. Figuring out mistakes after wrong answer in test series during preparation and correcting your bias afterwards.

Note:

- 1 All are supplement of each other, not an alternative.
- 2 Fundamentally different, but trying to achieve the same goal.

Machine learning: Type 3

17/ 44

Vishal Kumar
Jaiswal

Introduction

Machine Learning

Fundamentals

Inductive learning

Supervised Learning

Unsupervised
Learning

Reinforcement
Learning

Survey
Method

Mathematical
Preliminaries

Learning
Algorithms

Entropy

Ensemble
Learning
Method

Reinforcement learning

Learn how to behave based on delayed feedback from the environment.

e.g. Figuring out mistakes after wrong answer in test series during preparation and correcting your bias afterwards.

Note:

- 1 All are supplement of each other, not an alternative.
- 2 Fundamentally different, but trying to achieve the same goal.

Reinforcement learning

Learn how to behave based on delayed feedback from the environment.

e.g. Figuring out mistakes after wrong answer in test series during preparation and correcting your bias afterwards.

Note:

- 1 All are supplement of each other, not an alternative.
- 2 Fundamentally different, but trying to achieve the same goal.

- Collect data for machine learning by conducting survey or storing sensor output data stream.
- Preprocess data to remove inconsistencies in it.
- Before believing a survey result, check:
 - ① How many people or situations are surveyed.
 - ② Who is surveyed.
 - ③ How survey is conducted.
- Train and test machine learning algorithm on different sets of data, otherwise overfitting will happen.



Bayes' Theorem

19/ 44

Vishal Kumar
Jaiswal

Introduction

Machine Learning

Fundamentals

Inductive learning

Supervised Learning

Unsupervised

Learning

Reinforcement

Learning

Survey

Method

Mathematical
Preliminaries

Learning
Algorithms

Entropy

Ensemble
Learning
Method

Conditional Probability

Probability of event obtained with additional information that some other event has already occurred.

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

Assumption: We are dealing with sequential events and the new information is used to revise the probability of previous events.



Bayes' Theorem

20/ 44

Vishal Kumar
Jaiswal

Introduction

Machine Learning

Fundamentals

Inductive learning

Supervised Learning

Unsupervised
Learning

Reinforcement
Learning

Survey
Method

Mathematical
Preliminaries

Learning
Algorithms

Entropy

Ensemble
Learning
Method

Prior probability

Initial probability value obtained before any additional information is provided.

Posterior Probability

The probability value that has been revised by using additional information that is obtained later.

Bayes' Theorem

21/ 44

Vishal Kumar
Jaiswal

Introduction

Machine Learning

Fundamentals

Inductive learning

Supervised Learning

Unsupervised

Learning

Reinforcement

Learning

Survey
Method

Mathematical
Preliminaries

Learning
Algorithms

Entropy

Ensemble
Learning
Method

Definition

The probability of event A , given that event B has subsequently occurred, is

$$P(A|B) = \frac{P(A) \cdot P(B|A)}{[P(A) \cdot P(B|A)] + [P(\bar{A}) \cdot P(B|\bar{A})]}$$



Bayes' Rule: An example

22/ 44

Vishal Kumar
Jaiswal

Introduction

Machine Learning

Fundamentals

Inductive learning

Supervised Learning

Unsupervised

Learning

Reinforcement

Learning

Survey
Method

Mathematical
Preliminaries

Learning
Algorithms

Entropy

Ensemble
Learning
Method

A specific cancer, C , incurs in 1% of population, $P(C)=0.01$
Pathological Test:

Probability of positive test result, if you have C ,
 $P(Pos|C)=90\%$ (Sensitivity)

Probability of negative test result, if you don't have C ,
 $P(Pos|C)=90\%$ (Specificity) So, find out

- Probability of having cancer if test positive



Bayes' Rule: An example

22/ 44

Vishal Kumar
Jaiswal

Introduction

Machine Learning

Fundamentals

Inductive learning

Supervised Learning

Unsupervised

Learning

Reinforcement

Learning

Survey
Method

Mathematical
Preliminaries

Learning
Algorithms

Entropy

Ensemble
Learning
Method

A specific cancer, C , incurs in 1% of population, $P(C)=0.01$
Pathological Test:

Probability of positive test result, if you have C ,
 $P(Pos|C)=90\%$ (Sensitivity)

Probability of negative test result, if you don't have C ,
 $P(Pos|C)=90\%$ (Specificity) So, find out

- Probability of having cancer if test positive

Prior Probabilities:

$$P(C) = 0.01$$

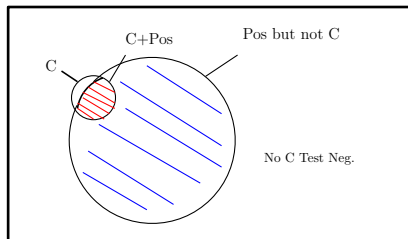
$$P(\neg C) = 0.99$$

$$P(Pos|C) = 0.9$$

$$P(Neg|\neg C) = 0.9$$

$$P(Pos|\neg C) = 0.1$$

All People





Prior Probabilities:

$$P(C) = 0.01 \quad P(\neg C) = 0.99$$

$$P(Pos|C) = 0.9$$

$$P(Neg|\neg C) = 0.9 \quad P(Pos|\neg C) = 0.1$$

Joint Probability:

$$P(C, Pos) = P(C) \times P(Pos|C) = 0.009$$

$$P(\neg C, Pos) = P(\neg C) \times P(Pos|\neg C) = 0.099$$

Normalize:

$$P(Pos) = P(\neg C, Pos) + P(C, Pos) = 0.108$$

$$\text{Posterior: } P(C|Pos) = \frac{P(C, Pos)}{P(Pos)} = 0.083$$

$$P(\neg C|Pos) = \frac{P(\neg C, Pos)}{P(Pos)} = 0.917$$

So, Total Probability=1



Prior Probabilities:

$$P(C) = 0.01 \quad P(\neg C) = 0.99$$

$$P(Pos|C) = 0.9$$

$$P(Neg|\neg C) = 0.9 \quad P(Pos|\neg C) = 0.1$$

Joint Probability:

$$P(C, Pos) = P(C) \times P(Pos|C) = 0.009$$

$$P(\neg C, Pos) = P(\neg C) \times P(Pos|\neg C) = 0.099$$

Normalize:

$$P(Pos) = P(\neg C, Pos) + P(C, Pos) = 0.108$$

$$\text{Posterior: } P(C|Pos) = \frac{P(C, Pos)}{P(Pos)} = 0.083$$

$$P(\neg C|Pos) = \frac{P(\neg C, Pos)}{P(Pos)} = 0.917$$

So, Total Probability=1

Prior Probabilities:

$$P(C) = 0.01 \quad P(\neg C) = 0.99$$

$$P(Pos|C) = 0.9$$

$$P(Neg|\neg C) = 0.9 \quad P(Pos|\neg C) = 0.1$$

Joint Probability:

$$P(C, Pos) = P(C) \times P(Pos|C) = 0.009$$

$$P(\neg C, Pos) = P(\neg C) \times P(Pos|\neg C) = 0.099$$

Normalize:

$$P(Pos) = P(\neg C, Pos) + P(C, Pos) = 0.108$$

$$\text{Posterior: } P(C|Pos) = \frac{P(C, Pos)}{P(Pos)} = 0.083$$

$$P(\neg C|Pos) = \frac{P(\neg C, Pos)}{P(Pos)} = 0.917$$

So, Total Probability=1

Prior Probabilities:

$$P(C) = 0.01 \quad P(\neg C) = 0.99$$

$$P(Pos|C) = 0.9$$

$$P(Neg|\neg C) = 0.9 \quad P(Pos|\neg C) = 0.1$$

Joint Probability:

$$P(C, Pos) = P(C) \times P(Pos|C) = 0.009$$

$$P(\neg C, Pos) = P(\neg C) \times P(Pos|\neg C) = 0.099$$

Normalize:

$$P(Pos) = P(\neg C, Pos) + P(C, Pos) = 0.108$$

$$\text{Posterior: } P(C|Pos) = \frac{P(C, Pos)}{P(Pos)} = 0.083$$

$$P(\neg C|Pos) = \frac{P(\neg C, Pos)}{P(Pos)} = 0.917$$

So, Total Probability=1

Prior Probabilities:

$$P(C) = 0.01 \quad P(\neg C) = 0.99$$

$$P(Pos|C) = 0.9$$

$$P(Neg|\neg C) = 0.9 \quad P(Pos|\neg C) = 0.1$$

Joint Probability:

$$P(C, Pos) = P(C) \times P(Pos|C) = 0.009$$

$$P(\neg C, Pos) = P(\neg C) \times P(Pos|\neg C) = 0.099$$

Normalize:

$$P(Pos) = P(\neg C, Pos) + P(C, Pos) = 0.108$$

$$\text{Posterior: } P(C|Pos) = \frac{P(C, Pos)}{P(Pos)} = 0.083$$

$$P(\neg C|Pos) = \frac{P(\neg C, Pos)}{P(Pos)} = 0.917$$

So, Total Probability=1

Prior Probabilities:

$$P(C) = 0.01 \quad P(\neg C) = 0.99$$

$$P(Pos|C) = 0.9$$

$$P(Neg|\neg C) = 0.9 \quad P(Pos|\neg C) = 0.1$$

Joint Probability:

$$P(C, Pos) = P(C) \times P(Pos|C) = 0.009$$

$$P(\neg C, Pos) = P(\neg C) \times P(Pos|\neg C) = 0.099$$

Normalize:

$$P(Pos) = P(\neg C, Pos) + P(C, Pos) = 0.108$$

$$\text{Posterior: } P(C|Pos) = \frac{P(C, Pos)}{P(Pos)} = 0.083$$

$$P(\neg C|Pos) = \frac{P(\neg C, Pos)}{P(Pos)} = 0.917$$

So, Total Probability=1

Prior Probabilities:

$$P(C) = 0.01 \quad P(\neg C) = 0.99$$

$$P(Pos|C) = 0.9$$

$$P(Neg|\neg C) = 0.9 \quad P(Pos|\neg C) = 0.1$$

Joint Probability:

$$P(C, Pos) = P(C) \times P(Pos|C) = 0.009$$

$$P(\neg C, Pos) = P(\neg C) \times P(Pos|\neg C) = 0.099$$

Normalize:

$$P(Pos) = P(\neg C, Pos) + P(C, Pos) = 0.108$$

$$\text{Posterior: } P(C|Pos) = \frac{P(C, Pos)}{P(Pos)} = 0.083$$

$$P(\neg C|Pos) = \frac{P(\neg C, Pos)}{P(Pos)} = 0.917$$

So, Total Probability=1

Prior Probabilities:

$$P(C) = 0.01 \quad P(\neg C) = 0.99$$

$$P(Pos|C) = 0.9$$

$$P(Neg|\neg C) = 0.9 \quad P(Pos|\neg C) = 0.1$$

Joint Probability:

$$P(C, Pos) = P(C) \times P(Pos|C) = 0.009$$

$$P(\neg C, Pos) = P(\neg C) \times P(Pos|\neg C) = 0.099$$

Normalize:

$$P(Pos) = P(\neg C, Pos) + P(C, Pos) = 0.108$$

$$\text{Posterior: } P(C|Pos) = \frac{P(C, Pos)}{P(Pos)} = 0.083$$

$$P(\neg C|Pos) = \frac{P(\neg C, Pos)}{P(Pos)} = 0.917$$

So, Total Probability=1

Prior Probabilities:

$$P(C) = 0.01 \quad P(\neg C) = 0.99$$

$$P(Pos|C) = 0.9$$

$$P(Neg|\neg C) = 0.9 \quad P(Pos|\neg C) = 0.1$$

Joint Probability:

$$P(C, Pos) = P(C) \times P(Pos|C) = 0.009$$

$$P(\neg C, Pos) = P(\neg C) \times P(Pos|\neg C) = 0.099$$

Normalize:

$$P(Pos) = P(\neg C, Pos) + P(C, Pos) = 0.108$$

$$\text{Posterior: } P(C|Pos) = \frac{P(C, Pos)}{P(Pos)} = 0.083$$

$$P(\neg C|Pos) = \frac{P(\neg C, Pos)}{P(Pos)} = 0.917$$

So, Total Probability=1

Prior Probabilities:

$$P(C) = 0.01 \quad P(\neg C) = 0.99$$

$$P(Pos|C) = 0.9$$

$$P(Neg|\neg C) = 0.9 \quad P(Pos|\neg C) = 0.1$$

Joint Probability:

$$P(C, Pos) = P(C) \times P(Pos|C) = 0.009$$

$$P(\neg C, Pos) = P(\neg C) \times P(Pos|\neg C) = 0.099$$

Normalize:

$$P(Pos) = P(\neg C, Pos) + P(C, Pos) = 0.108$$

$$\text{Posterior: } P(C|Pos) = \frac{P(C, Pos)}{P(Pos)} = 0.083$$

$$P(\neg C|Pos) = \frac{P(\neg C, Pos)}{P(Pos)} = 0.917$$

So, Total Probability=1

Prior Probabilities:

$$P(C) = 0.01 \quad P(\neg C) = 0.99$$

$$P(Pos|C) = 0.9$$

$$P(Neg|\neg C) = 0.9 \quad P(Pos|\neg C) = 0.1$$

Joint Probability:

$$P(C, Pos) = P(C) \times P(Pos|C) = 0.009$$

$$P(\neg C, Pos) = P(\neg C) \times P(Pos|\neg C) = 0.099$$

Normalize:

$$P(Pos) = P(\neg C, Pos) + P(C, Pos) = 0.108$$

$$\text{Posterior: } P(C|Pos) = \frac{P(C, Pos)}{P(Pos)} = 0.083$$

$$P(\neg C|Pos) = \frac{P(\neg C, Pos)}{P(Pos)} = 0.917$$

So, Total Probability=1

Prior Probabilities:

$$P(C) = 0.01 \quad P(\neg C) = 0.99$$

$$P(Pos|C) = 0.9$$

$$P(Neg|\neg C) = 0.9 \quad P(Pos|\neg C) = 0.1$$

Joint Probability:

$$P(C, Pos) = P(C) \times P(Pos|C) = 0.009$$

$$P(\neg C, Pos) = P(\neg C) \times P(Pos|\neg C) = 0.099$$

Normalize:

$$P(Pos) = P(\neg C, Pos) + P(C, Pos) = 0.108$$

$$\text{Posterior: } P(C|Pos) = \frac{P(C, Pos)}{P(Pos)} = 0.083$$

$$P(\neg C|Pos) = \frac{P(\neg C, Pos)}{P(Pos)} = 0.917$$

So, Total Probability=1



Naive Bayes Method

25/ 44

Vishal Kumar
Jaiswal

Introduction

Machine Learning

Fundamentals

Inductive learning

Supervised Learning

Unsupervised

Learning

Reinforcement

Learning

Survey
Method

Mathematical
Preliminaries

Learning
Algorithms

Entropy

Ensemble
Learning
Method

- Supervised learning algorithm based on Bayes' theorem.
- The "naive" assumption of independence between every pair of features.
- works quite well for many real-world applications. e.g. document classification and spam filtering.
- extremely fast
- decent classifier, but bad estimator.

Naive Bayes Method

25/ 44

Vishal Kumar
Jaiswal

Introduction

Machine Learning

Fundamentals

Inductive learning

Supervised Learning

Unsupervised

Learning

Reinforcement
Learning

Survey
Method

Mathematical
Preliminaries

Learning
Algorithms

Entropy

Ensemble
Learning
Method

- Supervised learning algorithm based on Bayes' theorem.
- The "naive" assumption of independence between every pair of features.
- works quite well for many real-world applications. e.g. document classification and spam filtering.
- extremely fast
- decent classifier, but bad estimator.

Naive Bayes Method

25/ 44

Vishal Kumar
Jaiswal

Introduction

Machine Learning

Fundamentals

Inductive learning

Supervised Learning

Unsupervised

Learning

Reinforcement
Learning

Survey
Method

Mathematical
Preliminaries

Learning
Algorithms

Entropy

Ensemble
Learning
Method

- Supervised learning algorithm based on Bayes' theorem.
- The "naive" assumption of independence between every pair of features.
- works quite well for many real-world applications. e.g. document classification and spam filtering.
- extremely fast
- decent classifier, but bad estimator.

Naive Bayes Method

25/ 44

Vishal Kumar
Jaiswal

Introduction

Machine Learning

Fundamentals

Inductive learning

Supervised Learning

Unsupervised

Learning

Reinforcement

Learning

Survey
Method

Mathematical
Preliminaries

Learning
Algorithms

Entropy

Ensemble
Learning
Method

- Supervised learning algorithm based on Bayes' theorem.
- The "naive" assumption of independence between every pair of features.
- works quite well for many real-world applications. e.g. document classification and spam filtering.
- extremely fast
- decent classifier, but bad estimator.



Naive Bayes Method

25/ 44

Vishal Kumar
Jaiswal

Introduction

Machine Learning

Fundamentals

Inductive learning

Supervised Learning

Unsupervised Learning

Reinforcement Learning

Survey
Method

Mathematical
Preliminaries

Learning
Algorithms

Entropy

Ensemble
Learning
Method

- Supervised learning algorithm based on Bayes' theorem.
- The "naive" assumption of independence between every pair of features.
- works quite well for many real-world applications. e.g. document classification and spam filtering.
- extremely fast
- decent classifier, but bad estimator.

Naive Bayes Method

25/ 44

Vishal Kumar
Jaiswal

Introduction

Machine Learning

Fundamentals

Inductive learning

Supervised Learning

Unsupervised

Learning

Reinforcement

Learning

Survey
Method

Mathematical
Preliminaries

Learning
Algorithms

Entropy

Ensemble
Learning
Method

- Supervised learning algorithm based on Bayes' theorem.
- The "naive" assumption of independence between every pair of features.
- works quite well for many real-world applications. e.g. document classification and spam filtering.
- extremely fast
- decent classifier, but bad estimator.

Gaussian Naive Bayes

26/ 44

Vishal Kumar
Jaiswal

Introduction

Machine Learning

Fundamentals

Inductive learning

Supervised Learning

Unsupervised
Learning

Reinforcement
Learning

Survey
Method

Mathematical
Preliminaries

Learning
Algorithms

Entropy

Ensemble
Learning
Method

Definition

The likelihood of features is:

$$P(x_i|y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp\left(-\frac{(x_i - \mu)^2}{2\sigma_y^2}\right)$$



उत्तिष्ठस्व आगत्य प्रत्यक्षं वदस्व
विद्यया विमुक्तो भवति
आर्य समाज, दिल्ली

27 / 44

Vishal Kumar
Jaiswal

Introduction

Machine Learning

Fundamentals

Inductive learning

Supervised Learning

Unsupervised

Learning

Reinforcement

Learning

Survey

Method

Mathematical

Preliminaries

Learning

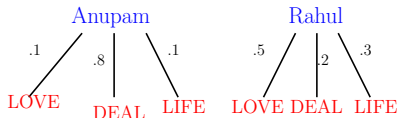
Algorithms

Entropy

Ensemble

Learning

Method



$$P(\text{Anupam}) = 0.5$$

$$P(\text{Rahul}) = 0.5$$

Guess: (Owner of email)

LOVE LIFE!

LIFE DEAL

Calculate

$$P(\text{"LIFE DEAL"} | \text{Anupam}) = 0.8 \times 0.1 \times 0.5$$

$$P(\text{"LIFE DEAL"} | \text{Rahul}) = 0.2 \times 0.3 \times 0.5$$



उत्तिष्ठत आगत्य प्रत्येकं वदतु विद्योपासकः
आत्मा विद्यया विद्यमानः ज्ञानं हि विद्योपासकः प्रथमं
आत्मैक्यं अनुभवति विद्यां तु द्वितीयं विद्यां, विद्यां

27 / 44

Vishal Kumar
Jaiswal

Introduction

Machine Learning

Fundamentals

Inductive learning

Supervised Learning

Unsupervised

Learning

Reinforcement

Learning

Survey

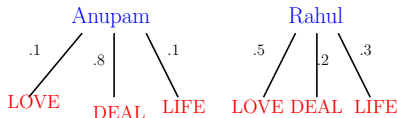
Method

Mathematical
Preliminaries

Learning
Algorithms

Entropy

Ensemble
Learning
Method



$$P(\text{Anupam}) = 0.5$$

$$P(\text{Rahul}) = 0.5$$

Guess: (Owner of email)

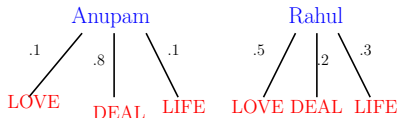
LOVE LIFE!

LIFE DEAL

Calculate

$$P(\text{"LIFE DEAL"} | \text{Anupam}) = 0.8 \times 0.1 \times 0.5$$

$$P(\text{"LIFE DEAL"} | \text{Rahul}) = 0.2 \times 0.3 \times 0.5$$



$$P(\text{Anupam}) = 0.5$$

$$P(\text{Rahul}) = 0.5$$

Guess: (Owner of email)

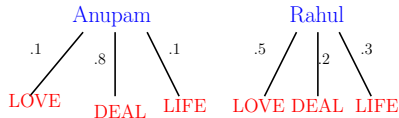
LOVE LIFE!

LIFE DEAL

Calculate

$$P(\text{"LIFE DEAL"} | \text{Anupam}) = 0.8 \times 0.1 \times 0.5$$

$$P(\text{"LIFE DEAL"} | \text{Rahul}) = 0.2 \times 0.3 \times 0.5$$



$$P(\text{Anupam}) = 0.5$$

$$P(\text{Rahul}) = 0.5$$

Guess: (Owner of email)

LOVE LIFE!

LIFE DEAL

Calculate

$$P(\text{"LIFE DEAL"} | \text{Anupam}) = 0.8 \times 0.1 \times 0.5$$

$$P(\text{"LIFE DEAL"} | \text{Rahul}) = 0.2 \times 0.3 \times 0.5$$

Support Vector Machine

28/ 44

Vishal Kumar
Jaiswal

Introduction

Machine Learning

Fundamentals

Inductive learning

Supervised Learning

Unsupervised
Learning

Reinforcement
Learning

Survey
Method

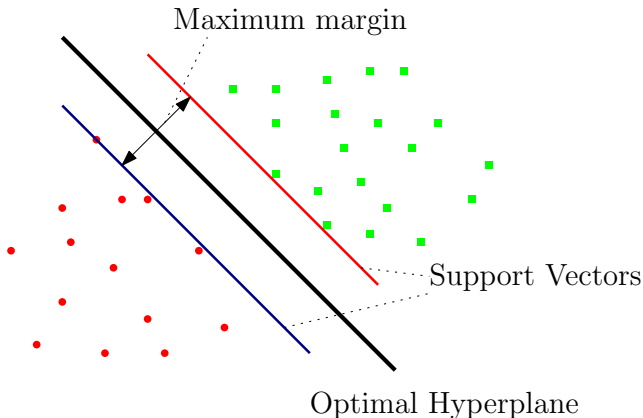
Mathematical
Preliminaries

Learning
Algorithms

Entropy

Ensemble
Learning
Method

- Supervised learning used for classification, regression analysis and outlier detection.
- Often used as a black box and for comparative study.
- Support Vectors are set of frontier datapoints of training dataset (may be linear).
- Effective in high dimensional space.



Pictorial representation of support vector machine



उत्तिष्ठस्व आगत्य पुण्यं वरदानं विद्योपायं
आयुः धनपुत्राणि विद्यायाः सागरः सौमित्रोऽसि
आसीत् अश्वमेधो विष्णुः सः शत्रुघ्नोऽसि, विष्णुः

30/ 44

Vishal Kumar
Jaiswal

Introduction

Machine Learning

Fundamentals

Inductive learning

Supervised Learning

Unsupervised
Learning

Reinforcement
Learning

Survey
Method

Mathematical
Preliminaries

Learning
Algorithms

Entropy

Ensemble
Learning
Method

```
# Support vector classification from sklearn module  
from sklearn.svm import SVC  
# Kernel: default: 'rbf', alternativess: 'linear','poly',  
clf=SVC(kernel="linear",gamma=1000.0,C=1)  
#Fit SVM model to according to the training data  
clf.fit(features_train,labels_train)  
# Perform classification on test set  
pred=clf.predict(features_test)
```

- Supervised algorithm
- Most popular, simple and extremely robust.
- Non-linear decision making with simple linear surfaces.

Disadvantage:

- over-complex trees that do not generalize well on datapoints(overfitting). Pruning is required.
- Learning optimal decision tree is NP-complete.



उत्प्रेषण आगत्य प्रत्येक वस्तुना विद्योपयुक्त
आवृत्तिप्रमाणेन विद्योपयुक्तं नाना विद्योपयुक्तं
आवृत्तिप्रमाणेन विद्योपयुक्तं नाना विद्योपयुक्तं

32/ 44

Vishal Kumar
Jaiswal

Introduction

Machine Learning

Fundamentals

Inductive learning

Supervised Learning

Unsupervised
Learning

Reinforcement
Learning

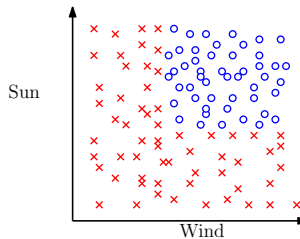
Survey
Method

Mathematical
Preliminaries

Learning
Algorithms

Entropy

Ensemble
Learning
Method



Weather pattern for Windsurf

Decision trees allows us to use multiple linear boundaries to separate two non-linearly separable regions.



उत्तिष्ठस्व आगत्य प्रत्यक्षं वदतां विद्योपायम्
आत्मकप्रज्ञायां विद्यमानं ज्ञानं नानुष्मन्तः शब्देन
आत्मनि ज्ञानमोक्षं विज्ञानं तत् तद्विज्ञानं ज्ञानम्, विज्ञानम्

32/ 44

Vishal Kumar
Jaiswal

Introduction

Machine Learning

Fundamentals

Inductive learning

Supervised Learning

Unsupervised
Learning

Reinforcement
Learning

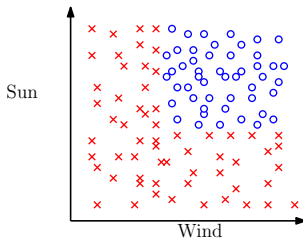
Survey
Method

Mathematical
Preliminaries

Learning
Algorithms

Entropy

Ensemble
Learning
Method



Weather pattern for Windsurf

Decision trees allows us to use multiple linear boundaries to seperate two non-linearly seperable regions.



उत्तिष्ठस्व आगत्य प्रपद्यस्व विद्यापरायणं
आत्मकमार्गं चतुर्वर्ण्यं ज्ञानं हि उपनिषदाः प्रथमं
आचार्यः श्रीमच्छंकराचार्यः

33/ 44

Vishal Kumar
Jaiswal

Introduction

Machine Learning

Fundamentals

Inductive learning

Supervised Learning

Unsupervised

Learning

Reinforcement

Learning

Survey

Method

Mathematical

Preliminaries

Learning

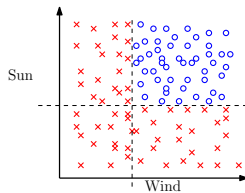
Algorithms

Entropy

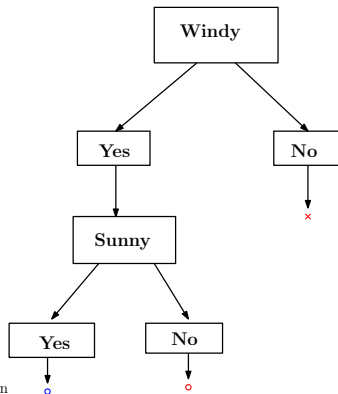
Ensemble

Learning

Method



Multiple linear decision boundary for weather pattern



Symbolic representation of decision tree for weather data

DecisionTreeClassifier is a class capable of performing multi-class classification on a dataset.

```
from sklearn import tree
```

```
X=[[0,0],[1,1]]
```

```
Y=[0,1]
```

```
clf=tree.DecisionTreeClassifier(min_samples_split=50)
```

```
clf=clf.fit(X,Y)
```

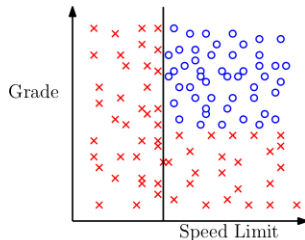
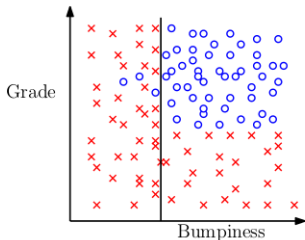
```
clf.predict([[2.,2.]])
```

Definition

Measure of impurity in a bunch of examples.

$$\text{Entropy}, H(P) = \sum -P_i \log_2 P_i$$

Controls how a decision tree decides to split the data, e.g.



Definition

The change in information entropy H from a prior state to a state that takes some information as given:

$$IG(T, a) = H(T) - H(T|a)$$

Decision trees tries to maximize information gain.

K Nearest Neighbors Classification Algorithm

37/ 44

Vishal Kumar
Jaiswal

Introduction

Machine Learning

Fundamentals

Inductive learning

Supervised Learning

Unsupervised

Learning

Reinforcement

Learning

Survey

Method

Mathematical
Preliminaries

Learning
Algorithms

Entropy

Ensemble
Learning
Method

- k closest training examples in the feature space are provided as input.
- Classification is done by a majority vote of neighbors.
- The object is assigned to the class most common among its k neighbors.

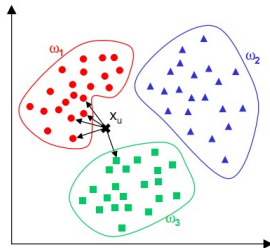


Illustration of K nearest neighbors

K Nearest Neighbors Classification Example

38/ 44

Vishal Kumar
Jaiswal

Introduction

Machine Learning

Fundamentals

Inductive learning

Supervised Learning

Unsupervised

Learning

Reinforcement

Learning

Survey

Method

Mathematical
Preliminaries

Learning
Algorithms

Entropy

Ensemble
Learning
Method

```
# Support vector classification from sklearn module  
from sklearn.svm import SVC  
# Kernel: default: 'rbf', alternativess: 'linear', 'poly',  
clf=SVC(kernel="linear",gamma=1000.0,C=1)  
#Fit SVM model to according to the training data  
clf.fit(features_train,labels_train)  
# Perform classification on test set  
pred=clf.predict(features_test)
```

Ensemble Learning Method

39/ 44

Vishal Kumar
Jaiswal

Introduction

Machine Learning

Fundamentals

Inductive learning

Supervised Learning

Unsupervised
Learning

Reinforcement
Learning

Survey
Method

Mathematical
Preliminaries

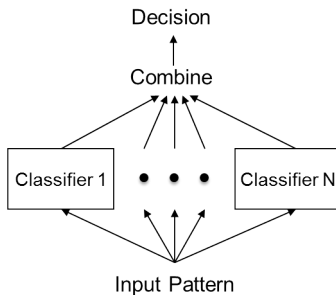
Learning
Algorithms

Entropy

Ensemble
Learning
Method

- Supervised Learning Methodology
- Uses multiple learning methods to obtain better performance.
- Prediction requires more computational overhead.
- Some unsupervised learning methods include consensus clustering and anomaly detection.

- Ensemble learning method
- Avoids overfitting of decision trees by building multiple deep decision trees
- Parameters include number of estimators.



Ensemble Learning Methodology

Adaptive Boosting, Adaboost

41/ 44

Vishal Kumar
Jaiswal

Introduction

Machine Learning

Fundamentals

Inductive learning

Supervised Learning

Unsupervised
Learning

Reinforcement
Learning

Survey
Method

Mathematical
Preliminaries

Learning
Algorithms

Entropy

Ensemble
Learning
Method

- Best known classifier
- Adaptive in the sense that subsequent weak learners are tweaked in favor of those instances misclassified by previous classifiers.
- Sensitive to noisy data and outliers
- AdaboostClassifier



उद्दिष्ट: आगमन, प्रगति, विकास, विद्योत्थान
मानव विकास, प्रगति, विकास, विद्योत्थान
मानव विकास, प्रगति, विकास, विद्योत्थान

42/ 44

Vishal Kumar
Jaiswal

Introduction

Machine Learning

Fundamentals

Inductive learning

Supervised Learning

Unsupervised
Learning

Reinforcement
Learning

Survey
Method

Mathematical
Preliminaries

Learning
Algorithms

Entropy

Ensemble
Learning
Method

- Numerical e.g. Salary info.
- Categorical e.g. Job Title
- Time Series e.g. time stamps on emails
- Text e.g. Contents of emails



Bayes' Theorem.



UD262: Machine Learning.



Ali Farhadi.

Cse446: Machine learning.
University of Washington.



Sebastian Thrun Katie Malone.

Ud102: Intro to machine learning.
Udacity.



पिंपरी चिंचवड एज्युकेशन ट्रस्ट
पिंपरी चिंचवड, महाराष्ट्र
१९८३

44 / 44

Vishal Kumar
Jaiswal

Introduction

Machine Learning

Fundamentals

Inductive learning

Supervised Learning

Unsupervised
Learning

Reinforcement
Learning

Survey
Method

Mathematical
Preliminaries

Learning
Algorithms

Entropy

Ensemble
Learning
Method

Thank You