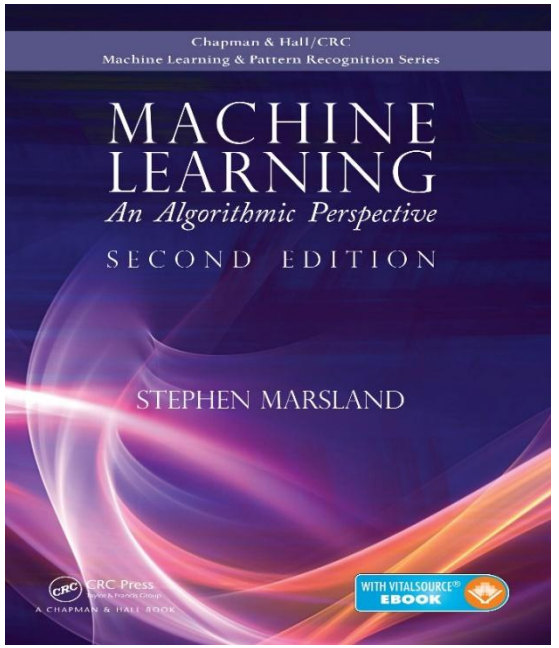
Unit 3 - Machine Learning Introduction, Cluster Analysis

- Machine Learning Introduction: Learning, Types of Machine Learning, Supervised Learning, The Machine Learning Process.

- Cluster Analysis: Basic concepts and methods: Cluster Analysis, Partitioning methods, Hierarchical Methods, Evaluation of clustering.

Text Books

- Stephen Marsland, “Machine Learning - An Algorithmic Perspective”, Second Edition, CRC Press - Taylor and Francis Group, 2015



Machine Learning

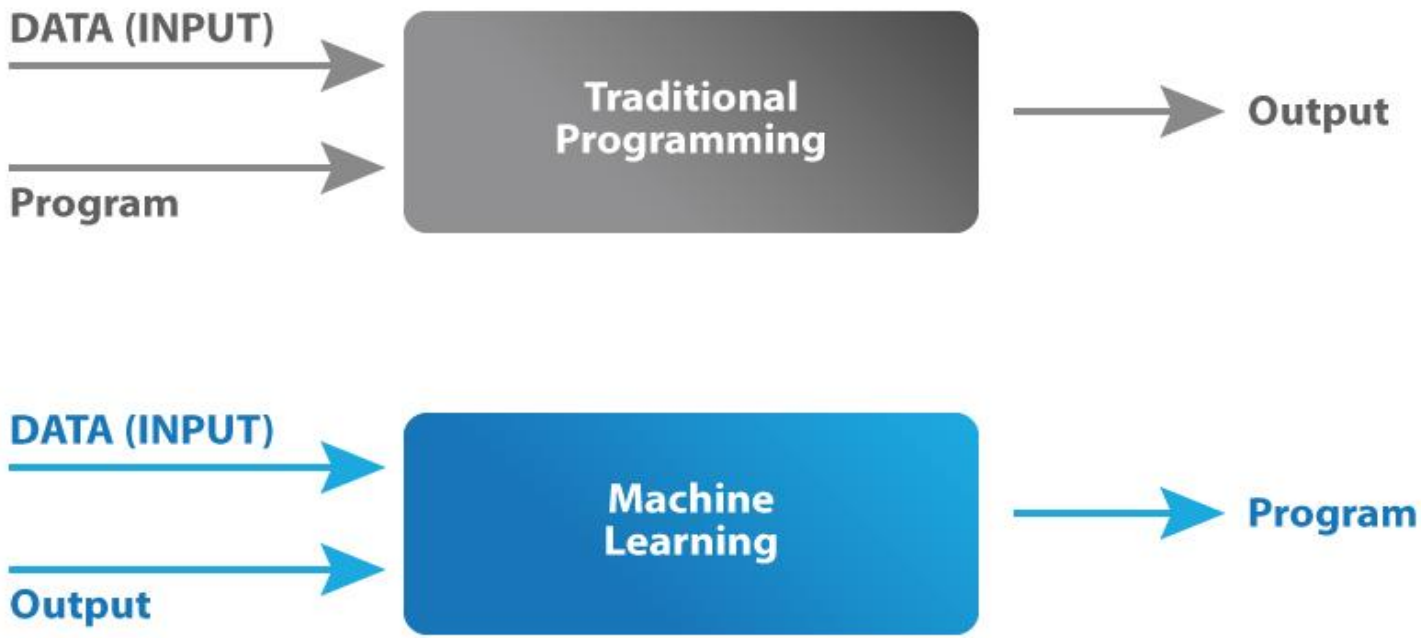
Topics

- **Machine Learning Introduction**
- Learning
- Types of Machine Learning
- Supervised Learning
- The Machine Learning Process

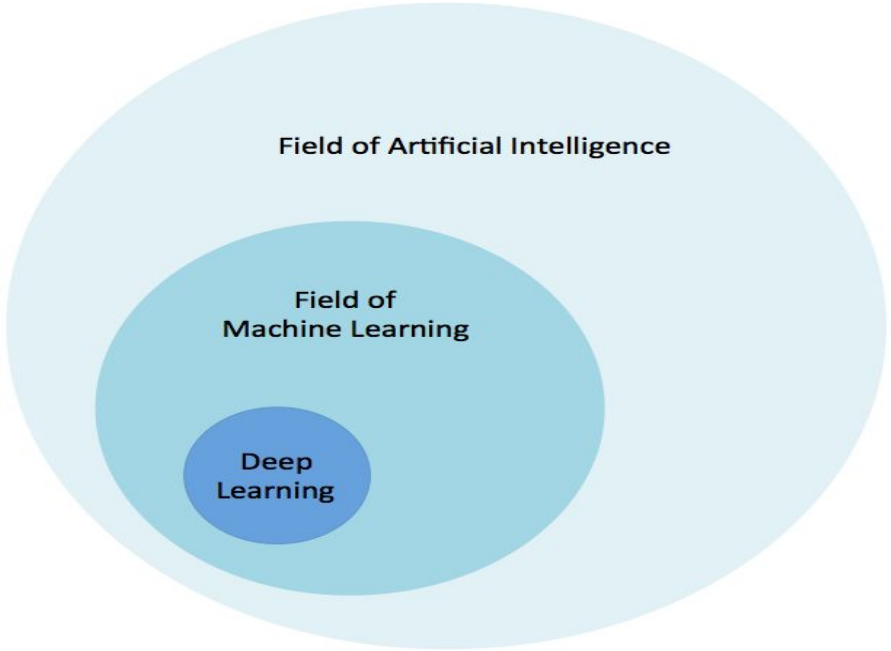
General Questions

- How our daily lives would be without Machine Learning?
- How our daily lives are with Machine Learning?
- What is the difference between Traditional Programming and Machine Learning?
- Why is Machine Learning important?

Traditional Programming/Machine Learning



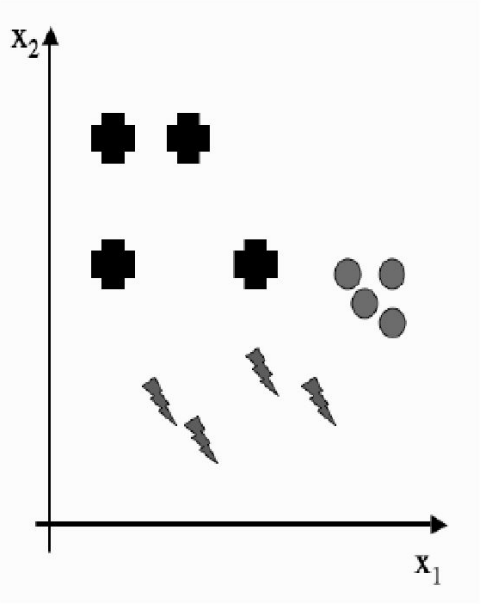
Introduction



Introduction

- Around the world, computers capture and store terabytes of data every day.
- The size and complexity of these datasets mean that humans are unable to extract useful information from them.unable to extract useful information from them.

x_1	x_2	Class
0.1	1	1
0.15	0.2	2
0.48	0.6	3
0.1	0.6	1
0.2	0.15	2
0.5	0.55	3
0.2	1	1
0.3	0.25	2
0.52	0.6	3
0.3	0.6	1
0.4	0.2	2
0.52	0.5	3



Introduction

- There are two things that we can do with this:
  - reduce the number of dimensions (until our simple brains can deal with the problem)
  - use computers, which don’t know that high-dimensional problems are difficult, and don’t get bored with looking at massive data files of numbers.

Introduction

- The two-dimensional projections of three-dimensional objects hides information.

Introduction

- This is one reason why machine learning is becoming so popular — the problems of our human limitations go away if we can make computers do the difficult work for us.
- There is one other thing that can help if the number of dimensions is not too much larger than three, which is to use glyphs that use other representations, such as size or colour of the datapoints to represent information about some other dimension
- But this does not help if the dataset has 100 dimensions in it.

Introduction

- Machine Learning are used in many of the software programs:
  - spam filters
  - voice recognition software
  - lots of computer games.
  - automatic number-plate recognition systems for petrol station security cameras and toll roads
  - vehicle stability systems
  - Decide whether a bank will give you a loan.

What is Machine Learning?

- **Machine Learning** is the field of study that gives computers the capability to **learn** without being explicitly programmed.
- **Improve** from experience.
- The core objective of machine learning is the **learning** and **inference**.
- The machine receives data as input, use an algorithm to formulate answers.
- Machine learning combines data with **statistical tools** to predict an output.
- The output is then used to make actionable insights.
- It is an application of Artificial Intelligence.

Why Machine Learning?

- Recent progress in algorithms and theory.
- Growing flood of online data.
- Huge memory and computational power for quicker processing.
- Industry demand.
- Less or no human intervention.
- Better decision making and predictions.

Relevance of Machine Learning in Industry

- Machine Learning Application in Healthcare Industry
  - Diagnostics
  - Breast Cancer/Diabetes/Covid-19 or any disease Prediction
  - Helps to identify cancerous tumors on mammograms.
  - Radiology and Radiotherapy
- Machine Learning Application in Agriculture Industry
  - Crop Management
  - Field Conditions Management
- Machine Learning Application in Education Industry
  - Personalized Learning

Relevance of Machine Learning in Industry

- Machine Learning Application in Transportation Industry
  - Traffic Regulation
  - Automated Driving
  - Vehicle quality assurance
- Machine Learning Application in Manufacturing Industry
  - Smart Maintenance
  - Better Product Development
  - Quality Improvement
  - Market Adaptation

Relevance of Machine Learning in Industry

- Credit card fraud detection
- Fake news detection
- Sentiment Analysis on Facebook
- Efficient model for searching and detecting duplicate questions in discussion forums
- Android malware classification in Android app
- Food and calories detection system.
- Mobile usage and health issues.
- Predicting whether the customer will churn or not in telecommunication industry.

Topics

- Machine Learning Introduction
- **Learning**
- Types of Machine Learning
- Supervised Learning
- The Machine Learning Process

Learning

- Our machines are learning from data, since data is what we have; terabytes of it, in some cases. However, it isn’t too large a step to put that into human behavioural terms, and talk about learning from experience.
- We are interested in the most fundamental parts of intelligence—learning and adapting—and how we can model them in a computer.
- This was the basis of most early Artificial Intelligence, and is sometimes known as symbolic processing because the computer manipulates symbols that reflect the environment.
- In contrast, machine learning methods are sometimes called subsymbolic because no symbols or symbolic manipulation are involved.

Learning

- Machine learning, then, is about making computers modify or adapt their actions (whether these actions are making predictions, or controlling a robot) so that these actions get more accurate, where accuracy is measured by how well the chosen actions reflect the correct ones.

Learning

- Imagine that you are playing Scrabble against a computer. You might beat it every time in the beginning, but after lots of games it starts beating you, until finally you never win. Either you are getting worse, or the computer is learning how to win at Scrabble. Having learnt to beat you, it can go on and use the same strategies against other players, so that it doesn't start from scratch with each new player;
- This is a form of generalisation.

Learning

- Machine Learning merges ideas from neuroscience and biology, statistics, mathematics, and physics, to make computers learn.
- Another thing that has driven the change in direction of machine learning research is data mining, which looks at the extraction of useful information from massive datasets and which requires efficient algorithms, putting more of the emphasis back onto computer science.

Learning

- The computational complexity of the machine learning methods will also be of interest to us since what we are producing is algorithms.
  - use some of the methods on very large datasets, so algorithms that have high degree polynomial complexity in the size of the dataset will be a problem.
  - The complexity is often broken into two parts: the complexity of training, and the complexity of applying the trained algorithm.
  - Training does not happen very often, and is not usually time critical, so it can take longer. However, we often want a decision about a test point quickly, and there are potentially lots of test points when an algorithm is in use.
  - This needs to have low computational cost.

What is Learning?

- Learning = Improving with experience at some task.
  - Improve over task T
  - With respect to performance measure P
  - Based on Experience E

When to Learn?

- Human expertise does not exist – navigating on Mars.
- Humans are unable to explain their expertise – Speech recognition.
- Solutions changes in time – Routing in networks.

Some Examples

- Example 1: Document/Text Classification.
- Example 2: Playing Chess.
- Example 3: Spam Emails/Non-Spam Emails.

Some Examples

- Example 4: A program to predict traffic patterns at a busy intersection (task T), you can run it through a machine learning algorithm with data about past traffic patterns (experience E) and, if it has successfully “learned”, it will then do better at predicting future traffic patterns (performance measure P).
- Example 5: A program that learns to play checkers might improve its performance as measured by its ability to win at the class of tasks involving playing checkers games, through experience obtained by playing games against itself.

What to Learn?

- Direct Learning (discriminative or model free) – learn a function that maps input and output.
- Indirect Learning (Generative or Model Learning) – learning a model of the domain, then use it to answer the questions about the domain.

Learning Problems

- A checkers learning problem:
  - Task T: playing checkers
  - Performance measure P: percent of games won against opponents
  - Training experience E: playing practice games against itself.

Student Task

1. Pick some learning task. Describe it by stating as precisely as possible the task, performance measure, and training experience.
    - **A handwriting recognition learning problem:**
    - **A robot driving learning problem:**

Topics

- Machine Learning Introduction
- Learning
- **Types of Machine Learning**
- Supervised Learning
- The Machine Learning Process

Types of Machine Learning

How does the computer know whether it is getting better or not, and how does it know how to improve?

- Supervised Learning
- Unsupervised Learning
- Reinforcement Learning
- Evolutionary Learning

Types of Machine Learning

- **Supervised learning** A training set of examples with the correct responses (targets) is provided and, based on this training set, the algorithm generalises to respond correctly to all possible inputs. This is also called learning from exemplars.
- **Unsupervised learning** Correct responses are not provided, but instead the algorithm tries to identify similarities between the inputs so that inputs that have something in common are categorized together. The statistical approach to unsupervised learning is known as density estimation.

Types of Machine Learning

- **Reinforcement learning** This is somewhere between supervised and unsupervised learning. The algorithm gets told when the answer is wrong, but does not get told how to correct it. It has to explore and try out different possibilities until it works out how to get the answer right. Reinforcement learning is sometime called learning with a critic because of this monitor that scores the answer, but does not suggest improvements.
- **Evolutionary learning** Biological evolution can be seen as a learning process: biological organisms adapt to improve their survival rates and chance of having offspring in their environment. We can model this in a computer, using an idea of fitness, which corresponds to a score for how good the current solution is.

Applications of Machine Learning

- **Virtual Personal Assistants** – Google Assistant, Alexa, Cortana, Siri.
- **Traffic Predictions** – Google Maps predicts traffic
  - Average Time taken on specific days at specific times on that route.
  - Real time location data of vehicles from Google maps application and sensors.
  - Other Map Services - Bing Maps, Maps.Me, HERE WeGo
- **Social Media Personalization**
- **Email Spam Filtering** – Spam Filters
  - Content Filters, Header Filters, General Black List Filters, Rules Based Filters etc.

Applications of Machine Learning

- **Online Fraud Detection – Fraud Risks**
  - Identity Theft
  - Fake Accounts
  - Man In The Middle Attacks
  - Feed Forward Neural Network – Genuine or Fraud Transactions
- **Stock Market Trading – Stock Market Indices – Nikkei**
  - LSTM NN – Predicting Trends
- **Assistive Medical Technology** – to diagnose diseases like Brain Tumors, Fetal Imaging, Cardiac Analysis
  - Personalized treatment, Drug Discovery, Clinical Research, Radiology

Applications of Machine Learning

- **Automatic Translation**
  - Sequence-to-Sequence Learning (same algo used in chatbots)
  - CNN - image recognition
  - Optical character recognition - to identify text.

Student Task

1. Pick some learning task. Describe it by stating as precisely as possible the task, performance measure, and training experience.

- **A handwriting recognition learning problem:**
  - Task T: recognizing and classifying handwritten words within images
  - Performance measure P: percent of words correctly classified
  - Training experience E: a database of handwritten words with given classifications

Student Task

1. Pick some learning task. Describe it by stating as precisely as possible the task, performance measure, and training experience.

- **A robot driving learning problem:**
  - Task T: driving on public four-lane highways using vision sensors
  - Performance measure P: average distance traveled before an error (as judged by human overseer)
  - Training experience E: a sequence of images and steering commands recorded while observing a human driver

Terminologies of Machine Learning

- **Model**
  A model is a **specific representation** learned from data by applying some machine learning algorithm. A model is also called **hypothesis**.
- **Feature**
  A feature is an individual measurable property of our data. A set of numeric features can be conveniently described by a **feature vector**. Feature vectors are fed as input to the model.

For example, in order to predict a fruit, there may be features like color, smell, taste, **etc.**

Note: Choosing informative, discriminating and independent features is a crucial step for effective algorithms. We generally employ a **feature extractor** to extract the relevant features from the raw data.

Terminologies of Machine Learning

- **Target (Label)**
  A target variable or label is the value to be predicted by our model. For the fruit example discussed in features section, the label with each set of input would be the name of the fruit like apple, orange, banana, etc.
- **Training**
  The idea is to give a set of inputs(features) and it's expected outputs(labels), so after training we will have a model (hypothesis) that will then map new data to one of the categories trained on.
- **Prediction**
  Once our model is ready, it can be fed a set of inputs to which it will provide a predicted output(label).

Machine Learning is Good

- **More Data > Better the Model > Higher Accuracy**

Student Task

- **Explore some interesting everyday examples around you where machines are learning and doing amazing jobs.**

Right Machine Learning Model Should be Used

- Algorithm to be used depends on:
  - The problem statement – Ex: Predict Future Stock Prices.
  - Size, quality and nature of the data.
  - Complexity of the algorithm.

- Machine Learning Algorithms are not types of Machine Learning.

Right Machine Learning Model Should be Used

- Classification
  - Falls under Supervised ML
  - Used when output is Yes/No or True/False
  - Spam Emails or not.
  - Algorithms are Decision Trees, Naïve Bayes, Logistic Regression, KNN, SVM, Adaboost, etc.
- Regression
  - Falls under Supervised ML
  - Used when a numerical value needs to be predicted like predict stock prices or predict the price of products based on the demand.
  - Algorithms are Linear Regression, Tree Based Regression, Multivariate Regression, etc.

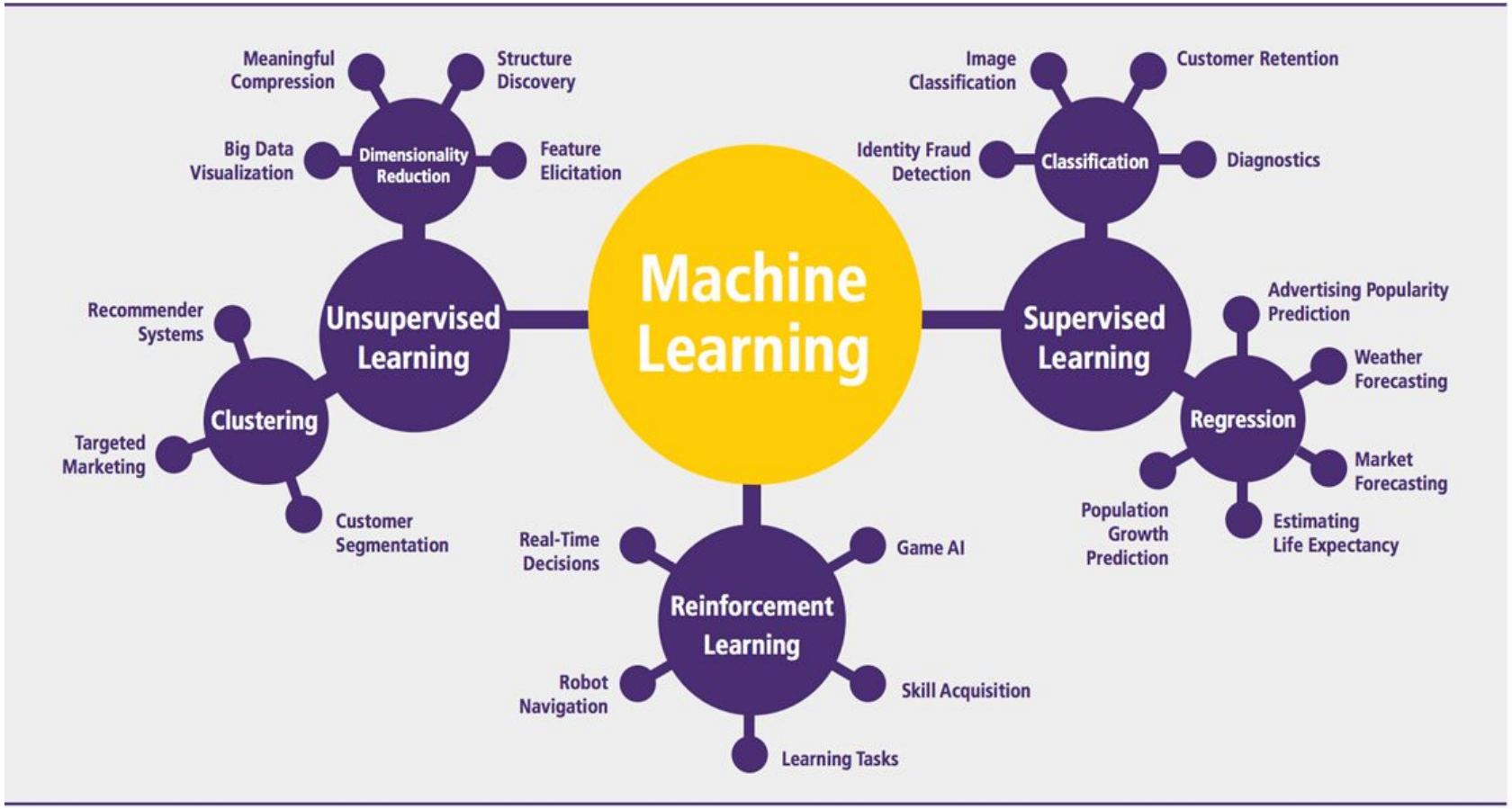
Right Machine Learning Model Should be Used

- Clustering
  - Falls under Unsupervised ML.
  - Used in search engines – provides best search results.
  - Used in Optimal Placement of Base Station.
  - Algorithms are k-means, hierarchical clustering, etc.
- Association
  - Falls under Unsupervised ML.
  - Used when data needs to be organized to find patterns in the product recommendation system like Amazon/Flipkart – Market Basket Analysis.
  - Algorithms are Apriori Algorithm, FP-Growth Algorithm, etc.

Challenges of Machine learning

- Data Representation.
- Machine Learning Algorithms Require Massive Stores of Training Data.
- Labeling Training Data Is a Tedious Process.
- Machines Cannot Explain Themselves.
- There is Bias in the Data.

Machine learning Algorithms and where they are used?



Given a Scenarios which type of learning is used and justify

- Filtering spam emails from email account – Classification Techniques.
- Predicting Stock Market Price, How much will it rain? – Regression Techniques.
- Facebook recognizes your friends picture from your album, Optimal Placement of Base Station in a given region – Clustering Techniques.
- Amazon/Flipkart checks and suggests which product suits us – Association Rule Mining.
- Netflix movie recommends some movie based on your past movie choice - Association Rule Mining.

Machine Learning Definitions

- Machine Learning is the field of study that gives computers the capability to **learn** without being explicitly programmed. **[Arthur Samuel 1959]**.
- A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance P on some task T, improves with experience E. **[Tom M Mitchell, 1997]**.
- “Machine learning (ML) is concerned with the design and development of algorithms and techniques that allow computers to “learn”. The major focus of ML research is to extract information from data automatically, by computational and statistical methods. It is thus closely related to data mining and statistics”. **[Svensson and Söderberg, 2008]**.

Advantages of Machine Learning

- Does not require knowledge engineers – everything we can handle with data.
- Scalable in constructing knowledge base.
- Adaptive to the changing conditions.
- Easy in migrating to new domains.

Limitations of Machine Learning

- Not useful on working on high dimensional data.
- Cannot solve problems like Image recognition and object recognition.
- Have to manually extract features and provide to the models, with huge dataset this is a tedious task.
- Cannot generate features on its own.

Why Deep Learning?

- Works on high dimensionality data.
- Huge memory and computational power for quicker processing.
- Industry demand.
- Better decision making and predictions.
- To focus on the right features by themselves.

Topics

- Machine Learning Introduction
- Learning
- Types of Machine Learning
- **Supervised Learning**
- The Machine Learning Process

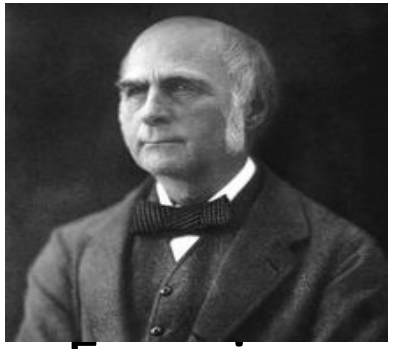
Supervised Learning

- **Classification**: Inputs are divided into two or more classes, and the learner must produce a model that assigns unseen inputs to one or more (multi-label classification) of these classes. This is typically tackled in a supervised way. Spam filtering is an example of classification, where the inputs are email (or other) messages and the classes are “spam” and “not spam”.
- **Regression**: It is also a supervised learning problem, but the outputs are continuous rather than discrete. For example, predicting the stock prices using historical data.

Linear Regression

- Statistical model which shows relation between two variables in a linear equation.
- Goal when using regression is to predict a numeric target value.
- One way to do this is to write out an equation for the target value with respect to the inputs.
- Linear regression means you can add up the inputs multiplied by some constants to get the output. $Y = mx + c$

The origins of Regression



- Regression was invented by the cousin of Charles Darwin, Francis Galton.
- Galton did his first regression in 1877 to estimate the size of pea seeds based on the size of their parents' seeds.
- He performed regression on a number of things, including the heights of humans. He noticed that if parents were above average in height, their children also tended to be above average but not as much as their parents. The heights of children were regressing toward a mean value.
- Galton noticed this behavior in a number of things he studied, and so the technique is called regression

English Meaning of Regression

- A return to a former or less developed state.
- "it is easy to blame unrest on economic regression“

Note: The English word having no relationship to predicting numeric values

Statistics Meaning of Regression

- A measure of the relation between the mean value of one variable (e.g. output) and corresponding values of other variables (e.g. time and cost).

What is Regression Analysis?

- Form of predictive modelling which investigates the relationship between dependent and independent variable.
- Drawing a line over a set of data points which mostly fits the overall shape of the data.
- Shows the changes in the dependent variable on the y-axis to the changes in the independent variable in the x-axis.

Uses of Regression Analysis

- Determining the strength of the predicators. Ex: sales and marketing, age and income.
- Forecasting an effect – impacts of changes (i.e. how much dependent variable changes with the change of one or more independent variable). Ex: how much addition sale income will I get for each $100 went in the marketing?
- Trend forecasting – predicts trends and future values. Ex: what is the price of the bitcoin in next 6months?

Linear Regression Algorithm

- The data is modelled using a straight line.
- Used with – continuous variable.
- Output/prediction – value of the variable.
- Accuracy and goodness of fit – measured by loss, R squared, Adjusted R squared, Root Mean Square etc.

Selection Criteria - Linear Regression

- Predicts continuous variables. Ex: Sales made for a day, temperature of a city.
- Not good for classification models.
- Data quality – each missing value can remove data points that could optimize the regression.
- Not computationally expensive as compared to decision trees, etc.
- Comprehensible and Transparent – simple mathematical notation and easy to understand.

Where is Linear Regression used?

- Evaluating trends and sales/forecast estimates – Ex: if a company has steadily increased every month from past few years then conduct linear regression analysis on the sales data with “monthly sales in y-axis and time on x-axis”. This gives the line that predicts the upper trends in the sales. Then use slope of the line to forecast sales in the future months.
- Analyzing the impact of the price changes – linear regression can be analyzed on the pricing based on the consumer behavior. Ex: Sole quantity as dependent and price as independent variable.
- Assessment of risks in financial services and insurance domain – Ex: no. of claims per customer against age.

Understanding Linear Regression

- Positive Regression Line
- Negative Regression Line
- Linear Regression Line – shows the relationship between independent and dependent variables.
- Plot actual data points on the graph
- Draw a linear regression line – find the error value between estimated point and actual point.
- The best fit line is the line with least error value or less distance between estimated point and actual point

Understanding Linear Regression

- Positive relationship y = mx + c speed v/s distance. Time constant
  - y is the distance travelled by the vehicle in a fixed amount of time – dependent variable.
  - x is the speed of the vehicle – independent variable.
  - m is the positive slope of the line.
  - c is the y-intercept of the line.
- Negative relationship y = -mx + c speed v/s time. Distance constant.
  - y is the time taken to travel a fixed distance.
  - x is the speed of the vehicle.
  - m is the negative slope of the line.
  - c is the y-intercept of the line.

Understanding Linear Regression

- x = (1,2,3,4,5), y = (3,4,2,4,5) – plot on graph.
- mean of x and y – Find and plot.
- The goal is to find best fit line using least square method.
- Find the regression line.
- Find ‘m’ ----- $m = (x – x') (y - y') / (x-x')^2$
- Find ‘c’
- Y = 0.4x + 2.4
- Find the predicted values or various value of x = 1,2,3,4,5.

Understanding Linear Regression

- The task is to find the distance between actual and predicted values. And reduce the distance or error value.
- Line with the least error will be the line of regression – Best fit line.
- It performs ‘n’ number of iterations and different value of ‘m’.
- For different values of ‘m’ it calculates the different equation of line.

Understanding Linear Regression

- As the value of ‘m’ changes – line is changing.
- For each iteration it will calculate the predicted value and then distance between actual and predicted values – Finds error value.
- For a particular ‘m’ – the least error value between actual and predicted values is selected as the best fit Regression line.
- Once you’ve found the regression weights, forecasting new values given a set of inputs is easy.

Goodness of Fit

- To check how good our model is performing.
- So use R-squared value.
- R-squared value is a statistical measure of how close the data are to the fitted regression line.
- Also known as coefficient of determination or coefficient of multiple determination .

Calculation of R^2

- Distance actual – mean vs Distance predicted – mean.
- $R^2 = (y_p - y')^2 / (y - y')^2$
- Calculate R^2
- R^2 = 0.3 which is not goodfit----actual and estimated points are far away.

Supervised Learning

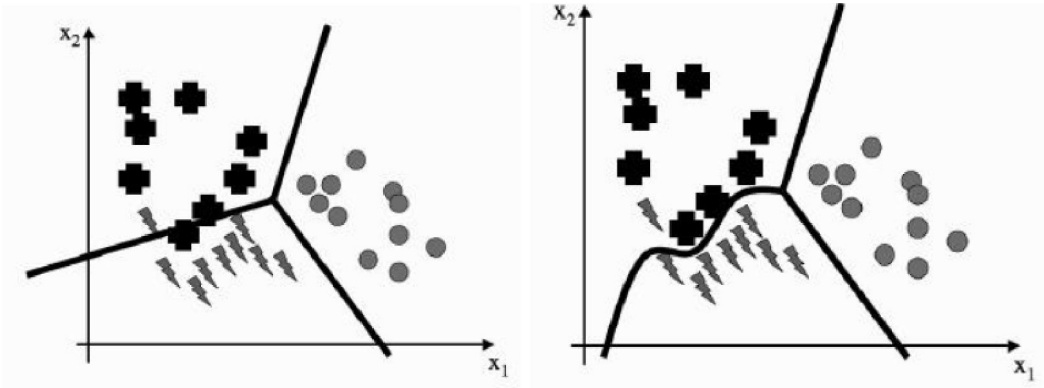
- Regression problem in statistics: fit a mathematical function describing a curve, so that the curve passes as close as possible to all of the datapoints.
- It is generally a problem of function approximation or interpolation, working out the value between values that we know.
- So one thing that our machine learning algorithms can do is interpolate between datapoints.
- This might not seem to be intelligent behaviour, or even very difficult in two dimensions, but it is rather harder in higher dimensional spaces.

Supervised Learning

- Classification—grouping examples into different classes. However, the algorithms are learning by our definition if they adapt so that their performance improves, and it is surprising how often real problems that we want to solve can be reduced to classification or regression problems.
- The classification problem consists of taking input vectors and deciding which of N classes they belong to, based on training from exemplars of each class.
- It is discrete—each example belongs to precisely one class, and the set of classes covers the whole possible output space.
- These two constraints are not necessarily realistic; sometimes examples might belong partially to two different classes - Fuzzy classifiers.
- In addition, there are many places where we might not be able to categorize every possible input.

Supervised Learning

- For example, consider a vending machine, where we use a neural network to learn to recognize all the different coins. We train the classifier to recognize all New Zealand coins, but what if a British coin is put into the machine?
- In that case, the classifier will identify it as the New Zealand coin that is closest to it in appearance, but this is not really what is wanted: rather, the classifier should identify that it is not one of the coins it was trained on. This is called novelty detection.



Supervised Learning

Coin Classifier:

- When the coin is pushed into the slot, the machine takes a few measurements of it.
- These could include the diameter, the weight, and possibly the shape, and are the features that will generate our input vector.
- In this case, our input vector will have three elements, each of which will be a number showing the measurement of that feature (choosing a number to represent the shape would involve an encoding, for example that 1=circle, 2=hexagon, etc.).

Supervised Learning

Coin Classifier:

- Of course, there are many other features that we could measure like the density of the material and its composition, or if it had a camera, we could take a photograph of the coin and feed that image into the classifier.
- The question of which features to choose is not always an easy one.
- We don’t want to use too many inputs, because that will make the training of the classifier take longer
- As the number of input dimensions grows, the number of datapoints required increases faster - curse of dimensionality

Supervised Learning

Coin Classifier:

- Make sure that we can reliably separate the classes based on those features.
- Find decision boundaries that can be used to separate out the different classes.
- Given the features that are used as inputs to the classifier, we need to identify some values of those features that will enable us to decide which class the current input is in.

Topics

- Machine Learning Introduction
- Learning
- Types of Machine Learning
- Supervised Learning
- **The Machine Learning Process**

The Machine Learning Process

- **Data Collection and Preparation -** Either the data has to be collected from scratch, or at the very least, assembled and prepared.
- In fact, if the problem is completely new, so that appropriate data can be chosen, then this process should be merged with the next step of feature selection, so that only the required data is collected.
- This can typically be done by assembling a reasonably small dataset with all of the features that you believe might be useful, and experimenting with it before choosing the best features and collecting and analysing the full dataset.

The Machine Learning Process

- **Data Collection and Preparation -** For supervised learning, target data is also needed, which can require the involvement of experts in the relevant field and significant investments of time.
- Finally, the quantity of data needs to be considered.
- Machine learning algorithms need significant amounts of data, preferably without too much noise, but with increased dataset size comes increased computational costs.
- If there is enough data without excessive computational overhead is generally impossible to predict.

The Machine Learning Process

- **Feature Selection** - For coin recognition. It consists of identifying the features that are most useful for the problem under examination.
- This invariably requires prior knowledge of the problem and the data.
- The features can be collected without significant expense or time, and that they are robust to noise and other corruption of the data that may arise in the collection process.

The Machine Learning Process

- **Algorithm Choice** - Given the dataset, the choice of an appropriate algorithm (or algorithms) in that the knowledge of the underlying principles of each algorithm and examples of their use is precisely what is required for this.

- **Parameter and Model Selection** - For many of the algorithms there are parameters that have to be set manually, or that require experimentation to identify appropriate values.

The Machine Learning Process

- **Training** - Given the dataset, algorithm, and parameters, training should be simply the use of computational resources in order to build a model of the data in order to predict the outputs on new data.

- **Evaluation** - Before a system can be deployed it needs to be tested and evaluated for accuracy on data that it was not trained on. This can often include a comparison with human experts in the field, and the selection of appropriate metrics for this comparison.

Student Task

- **Explore some interesting everyday examples around you where machines are learning and doing amazing jobs.**

- **Answers in next 3 slides**

Some Examples of Machine Learning Problems

- “Is this cancer?”, “What is the market value of this house?”, “Which of these people are good friends with each other?”, “Will this rocket engine explode on take off?”, “Will this person like this movie?”, “Who is this?”, “What did you say?”, and “How do you fly this thing?”.
- All of these problems are excellent targets for an ML project, and in fact ML has been applied to each of them with great success.

Machines are learning and doing amazing jobs…

- Machines learning from medical records which treatments are most effective for new diseases
- Learning from experience to optimize house energy costs based on the particular usage patterns of their occupants
- Personal software assistants learning the evolving interests of their users in order to highlight especially relevant stories from the online morning newspaper.

Machines are learning and doing amazing jobs…

- In the field known as data mining, machine learning algorithms are being used routinely to discover valuable knowledge from large commercial databases containing equipment maintenance records, loan applications, financial transactions, medical records, etc.
- Programs have been developed that successfully learn to:
  - Recognize spoken words
  - Predict recovery rates of pneumonia patients
  - Detect fraudulent use of credit cards
  - Drive autonomous vehicles on public highways
  - Play games such as backgammon at levels approaching the performance of human world champions.

Machine Learning Frameworks

- TensorFlow
- Theano
- Scikit-learn
- Caffe
- Amazon Machine Learning
- Google Cloud ML Engine
- Azure ML studio
- Spark ML lib

Machine Learning Resources

- Kaggle Competitions - https://www.kaggle.com/competitions
- Check UCI dataset – which has various datasets from different domains.
- Relevant journals include:
  - Machine Learning
  - Neural Computation
  - Neural Networks
  - Journal of the American Statistical Association
  - IEEE Transactions on Pattern Analysis and Machine Intelligence.

Machine Learning Resources

- Annual Conferences of machine learning:
  - International Conference on Machine Learning
  - Neural Information Processing Systems
  - Conference on Computational Learning Theory
  - International Conference on Genetic Algorithms
  - International Conference on Knowledge Discovery and Data Mining
  - European Conference on Machine Learning