Applications of ML in Industries

Goutam Datta
Dept. of Informatics, School of Computer Science, UPES,Dehradun

A Brief Introduction to ML

- Paradigms,
- Knowledge Representation,
- Data Acquisition, Data Pre-Processing,
- Feature Extraction and Processing
- Feature Ranking and Selection, Feature Reduction,
- Model Learning,
- Evaluation and Deployment

Paradigms

- Classical to Modern AI
- Classical AI :

  Symbolic

  Rule Based

  Difficult to learn by itself

  mainly Heuristic based approach

  Reasoning based

  Expert systems

  Knowledge Representation

Modern AI

- Ability to learn by itself
- Data driven based
- Neural Network based
- **Fuzzy logic based**
- Neuro Fuzzy based
- Genetic algorithms based
- Optimization based
- Overall soft computing based

Knowledge Representation

- Predicate logic
- Propositional logic
- Semantic Net
- Conceptual Dependency
- Frames

Data Acquisition, Preprocessing,Feature Extraction

- ML demands two primary things : Data and model .
- During data acquisition features are to be maintained/selected adequately.
- With proper features, model can be trained properly.
- data gathered .
- When data available it has to be prepared with the help of preprocessing : Preprocessing involves multiple transformation operations such as filtering, sampling, scaling, reducing, editing columns, finding missing values etc.

Model : Training and Tuning

- First figure out type of analysis. Classification analysis (spam vs non-spam), a clustering analysis (automatic classification), or a regression analysis (for prediction and forecasting). compare the results of multiple models to know which one is the most efficient for your dataset. The logic here is to:

1. Select a model with default value
2. Train the model
3. Score the model
4. Evaluate the model (to figure out which model is the most efficient)
5. Sweep the model (to figure out the best configuration for your model)
6. Evaluate the model
7. Save the trained model (to be used in production)

ML in Banking and Securities: Major Areas of Banking with Potential AI intervention

- Anti Money Laundering(AML)
- Chatbot
- Fraud Detection
- Algorithmic trading and digitization: Executed by institutional traders and large Brokers.

Widely used ML Algorithm in banking and security

- Supervised ML algorithm
- Unsupervised ML algorithm
- Reinforcement ML algorithm

Domains

- Fraud and risk management
- Customer services
- Financial trading and securities
- Credit assessment
- Portfolio management

Content Based Image Retrieval System

- ML algorithm/Soft Computing based approach to reduce semantic gap.

Role of ML : Challenges of banking sector and securities

- Fraud and Risk Management : ML based solutions and predictive analytics are assisting in examination of real time transaction to identify suspicious and fraudulent operations.
- Customer Services : ML powered AI applications helps analyzing customer behavior and requirements.
- Financial trading and securities : AI based applications are assisting banks in effectively handling foreign exchange transaction and liquidity management.
- Credit assessment : ML based applications along with big data analytics are viable solution to assess the credit worthiness of the customer in case of loan disbursement.
- Portfolio management :AI and ML based technological ecosystems are helping banks in real time, smarter decisions to ensure appropriate investment plans.

Widely used ML algorithm in banking and security

- Supervised ML algorithms
- Unsupervised ML algorithms
- Reinforcement algorithms

Portfolio Management System

- Portfolio : It can be defined as financial assets like stocks, bonds, shares,mutual funds, debt instruments etc.
- In order to maintain the risk in various asset pools of investment, a portfolio is planned.

Management : In order to achieve its pre described objectives with well defined policies, a management can be defined as a firm which coordinates the activities of that firm/enterprise.

Objectives of Portfolio Management

- Security of Principal Investment
- Consistency of Returns
- Capital Growth
- Marketability
- Liquidity
- Diversification of Portfolio
- Favorable Tax Status.

Deep Learning For Customer Services

- Chatbots : The ability of chatbots is used as effective self service solution that interacts with the customer service representative and resolve their queries.
- ML allows chatbots to analyze and know when to utilize answers,collect customer details and turn over the interaction to the individual customer support agents.
- ML helps chatbots to know when to utilize unique answers, when to collect required user information and when to send a discussion to a person agent.

Recurrent Neural Network

- The main problem with feed forward neural network is it only focuses on the current value/context.
- This value is not retained in the next time stamp.
- Sometimes in the real world situation we have to retain the previous value and based on the previous and current input/context we have to decide the output.
- For example, in case of sentences what will be the next word, that can be decided with the help of the context of the previous words.

Recurrent Neural Network Contd.

- Hence in language modelling we require long term temporal dependencies to decide the next word in the sentences.
- RNN model helps in such type of situation. Long term temporal dependencies can not be achieved with feed forward neural network.
- A simple model of recurrent neural network uses tanh or sigmoid as an activation function.
- Recurrent neural network also suffers from vanishing gradient and exploding gradient problem during the training process

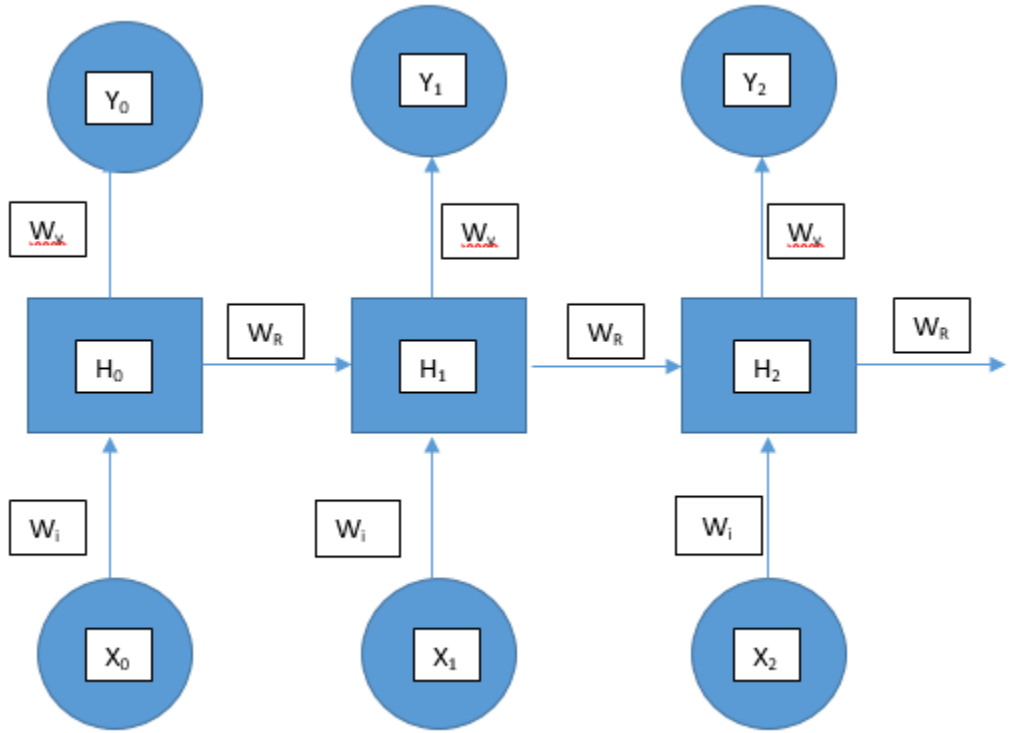
RNN contd.

- When training a deep neural network with gradient based learning and backpropagation, we find the partial derivatives by traversing the network from the the final layer to the initial layer.
- Using the chain rule, layers that are deeper into the network go through continuous matrix multiplications in order to compute their derivatives.

RNN Contd.

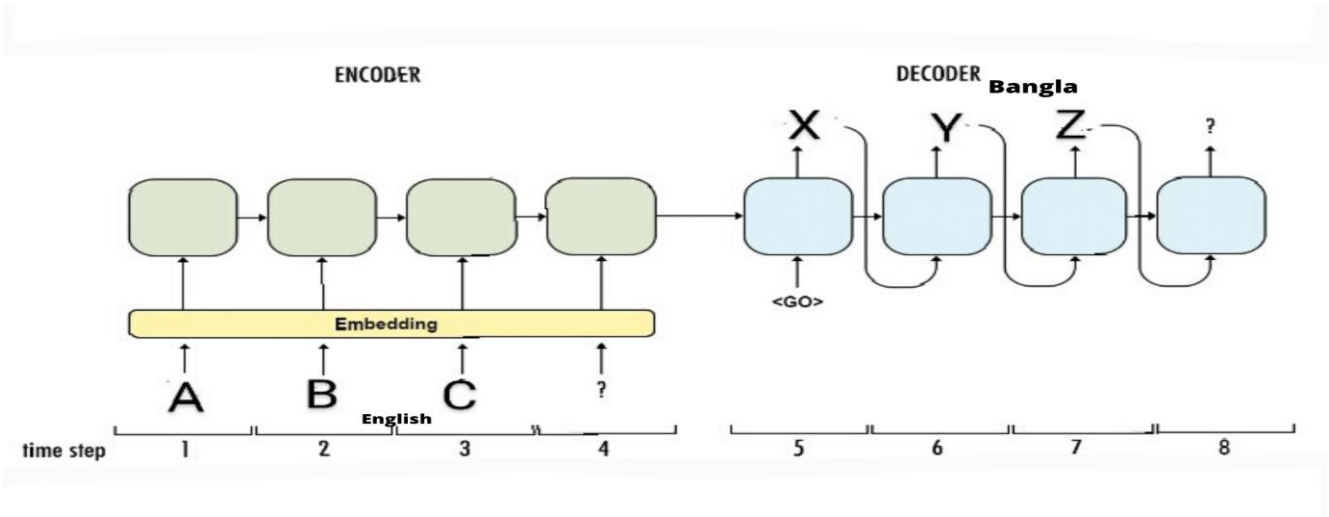
- In a network of *n* hidden layers, *n* derivatives will be multiplied together.
- If the derivatives are large then the gradient will increase exponentially as we propagate down the model until they eventually explode, and this is what we call the problem of ***exploding gradient***.
- Alternatively, if the derivatives are small then the gradient will decrease exponentially as we propagate through the model until it eventually vanishes, and this is the ***vanishing gradient*** problem.

RNN Contd.

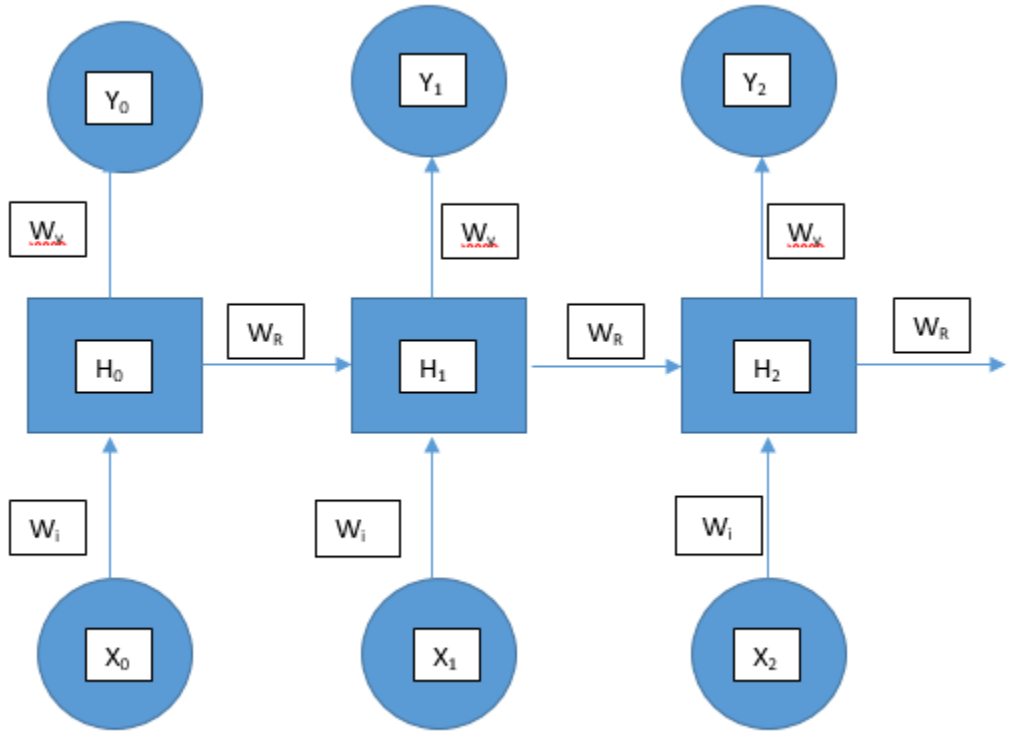


Chatbot : Deep learning approach

- The widely using deep learning model in chatbot is sequence to sequence model.
- Sentiment Analysis



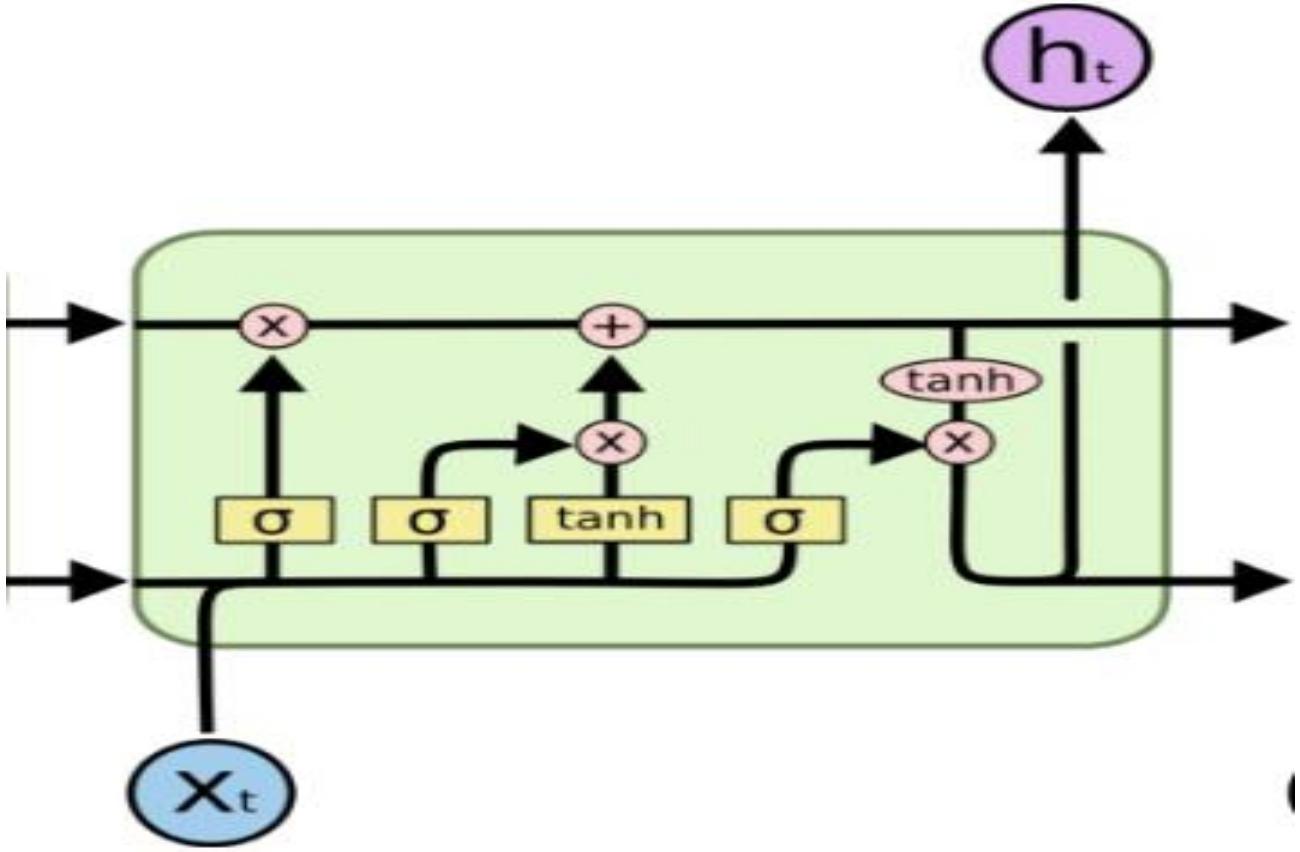
Recurrent Neural Network



Long Short Term Memory(LSTM)

- The long term dependency problem can be solved with LSTM models.
- LSTM cells are operated with the help of real number parameter called gates.
- The input gate parameter helps to decide how much new input is required to change the memory state.
- The forget gate parameter helps to decide how much previous value to retain in-memory state.
- And the output gate parameter controls how strongly the current memory state is to pass into the next layer.

LSTM Architecture



Fraud Detection : Few Supervised Learning Algorithms.(IBM –Text Material)

- Random Forest :
- Support Vector Machines:
- K-Nearest Neighbors(KNN):

Random Forest :

- Random forest is an ensemble of decision trees, i.e. develops a group of decision trees to classify data objects.
- Advantages of Random Forest :

❖The algorithm is simple and exhibits higher operating speed.

❖It can be used with variety of data including credit and debit card numbers, dates,IP(Internet Protocol)addresses, Postal Index Numbers etc.

❖They are efficient predictors that can perform well even with datasets having considerable missing records.

Random Forest

- Disadvantages of Random Forest :

❖Prediction accuracy of random forest reduces considerably when training data contains more of normal transaction instances and very less proportion of fraud cases.

❖Overfitting problem – a case where a model considers noise and fluctuations in the training data set as learning instances.

Support Vector Machine(SVM)

- Algorithm works based on dividing given data set in to two distinct categories with well defined gap.
- Several hyper planes are made to identify which plane clearly divides the data set.
- Advantages of SVM :

❖SVM are more suitable to work with complex multidimensional systems.

❖The challenge of overfitting is avoided, which is generally experienced in case of random forest.

❖It is simple and yet effective in terms of accuracy but some time slower in its operation and demands high end computing ecosystem.

K-Nearest neighbor

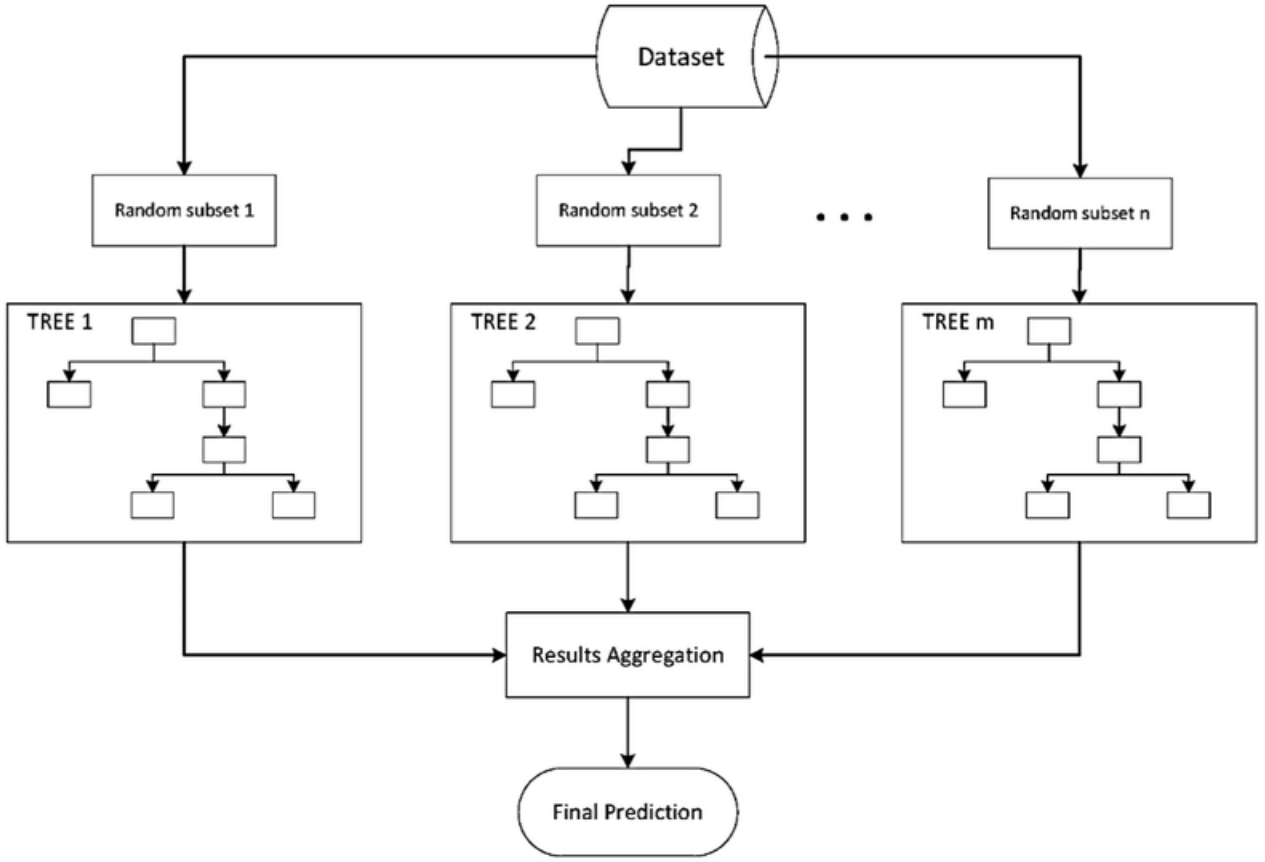
- KNN algorithm works with an assumption that similar kind of objects/instances exist in close vicinity or distance or proximity.
- With the given data set, the algorithm compares the distance between current data and the newly available data instance to identify its neighborhood or proximity.
- Advantages of KNN:

❖Implementation of algorithm is easy.

❖It can be effectively used for classification, regression etc.

❖Drawback : Speed of the operation reduces significantly with increase number of data set to be compared to decide the close proximity.

Ensemble method



Ensemble contd.

- The goal of any machine learning problem is to find a single model that will best predict our wanted outcome. Rather than making one model and hoping this model is the best/most accurate predictor we can make, ensemble methods take a myriad of models into account, and average those models to produce one final model. It is important to note that Decision Trees are not the only form of ensemble methods, just the most popular and relevant in DataScience today.

Case Study : Sentiment Analysis for student's feedback using ML

- The study focuses on symbolic analysis for online reviews from participants.
- Identify the feeling present in text documents based on polrity- positive, negative and neutral using the approach of machine learning, Lexicon based and Hybrid.
- Classifiers are used to measure the performance and accuracy.

Machine Learning Techniques for Customer Sentiment Analysis

- Application of NLP
- Company should also have knowledge on the following aspects

❖What is the response the marketing campaign getting?

❖Which different age group respond?

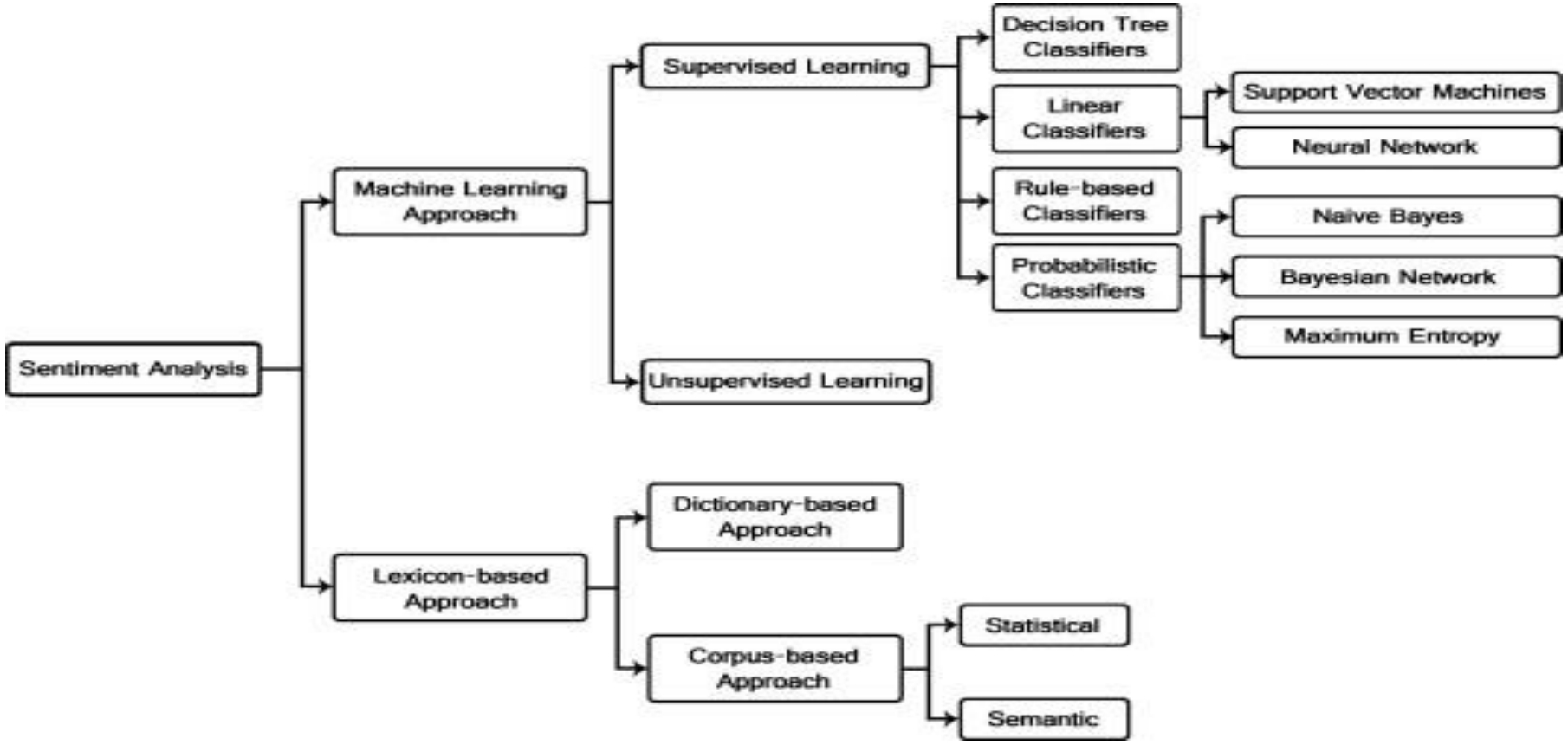
❖What’s the difference in men and women responding? etc. etc.

Sentiment Analysis for student's Feedback: Work Flow

Steps :

- Student's feedback : Awesome lecture, good delivery, boring
- Preprocessing : Tokenization, Stop word removal, POS tagging
- Feature Extraction :
- Classification : Random forest
- Evaluation : Model evaluation
- Accuracy, precision, visualization etc.

Sentiment Analysis for student's feedback using ML



Machine Learning Based Sentiment Analysis

- Machine learning based approaches of sentiment analysis learn a predictive model using the provided training dataset and evaluate the performance of the learned model on the test dataset.
- It can be further classified into supervised learning and unsupervised learning methods.

Lexicon based Method :

- Lexicon based approach of sentiment analysis makes use of a sentiment lexicon to determine the polarity of a given textual content.
- A lexicon or dictionary represents a list of words with associated sentiment polarity. The lexicon can be constructed either manually or automatically.
- Hu and Liu [7]utilized an online lexical resource WordNet to predict the semantic orientation of an opinion word.
- Taboada et al. [8]proposed another lexicon-based approach that determines the polarity of a word by using the dictionaries constructed.

Hybrid Approach :

- Combination of Lexicon based and Machine Learning based approach
- An opinion lexicon was used to label training dataset with sentiment polarities. The labeled dataset was then used to train a binary classifier to predict sentiment polarity on the evaluation dataset.

Reference

- “Learning sentiment from students’ feedback for real-time interventions in classrooms”, Nabeela Altrabsheh, Mihaela Cocea, and Sanaz Fallahkhair

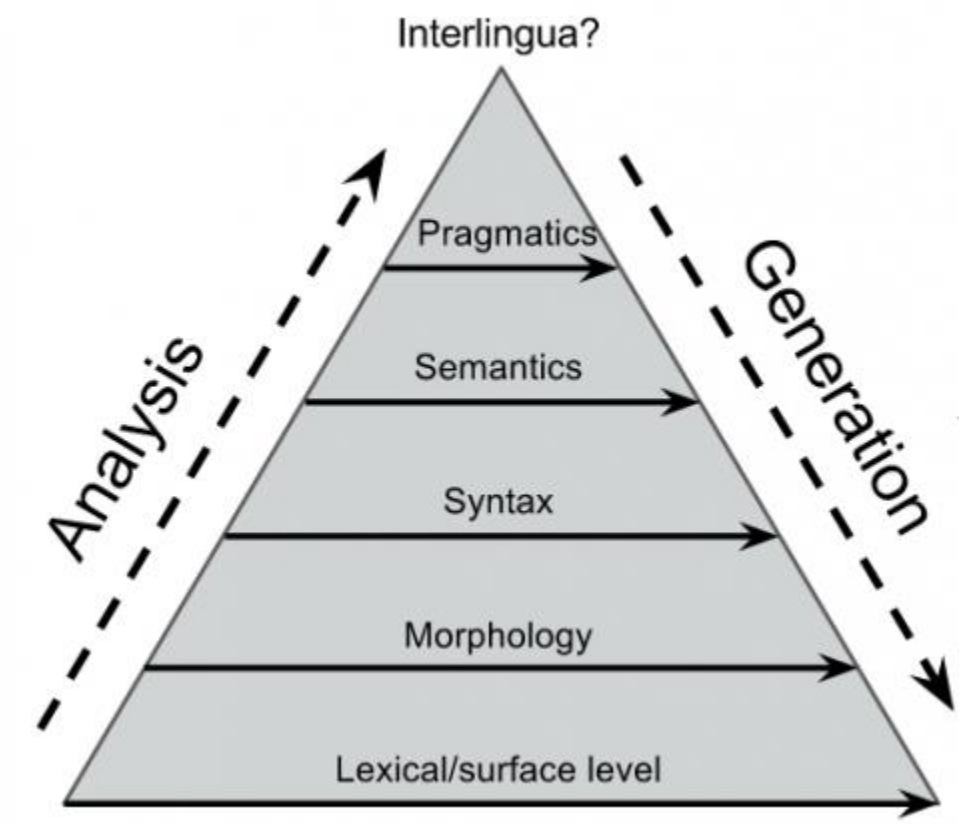
Credit Default Prediction

- https://www.kaggle.com/uciml/default-of-credit-card-clients-dataset

Machine Translation in a Nutshell

- Machine translation(MT) as its name implies translate text written in one language to text written in other language.

Vacuous Triangle



Linguistic Aspects of MT system

- Lexical transfer is like a dictionary where we can get the meaning of words.
- Morphology means the structure of the word, for example, goes is recognized as inflected third person singular number present form of the verb go.
- Syntax is the grammar of the language and semantics is the meaning

of the word/sentence and so on.

MT Contd.: MT steps…..

- **Preprocessing :** Process the text by removing punctuation,
- Converting text to lower and/or true casing etc.
- Pre processing could be tokenization, padding etc.
- Tokenization is the process where we break the entire phrase, sentence into small units called words or tokens.
- Padding is done in translation task to find end of sentence.
- Padding is context/task dependent i.e. padding may have different purpose for different NLP application.
- To make all the sequence word size of equal length wherever any word size is smaller than the fixed maximum size, padding is used to make that equal to that fixed maximum size.

MT Contd…..

- **Modeling:** There are three important steps involved in modeling. They are building a model, then train the model and finally testing the model.
- **Prediction:** We have to translate English to Other language using our model and then have to compare this translation with ground truth and finally can evaluate on the basis of different automatic metric.
- **Iteration:** Iterating on the model by different hyper parameter adjustments and also on different architectures the model will be tried to find better accuracy in terms of translation.

MT Contd. : Statistical MT System

- In this approach machine learns from the data automatically with the help of statistics.
- It is actually on the basis of machine learning approach where it tries to map a proper function for source to target language.
- Unlike machine learning, the function in SMT is non deterministic.
- In machine learning, classification is actually considered as mapping from many to one.
- But in SMT we can map from one source language to multiple possible representation of target language i.e. one to many and hence, considered non-deterministic

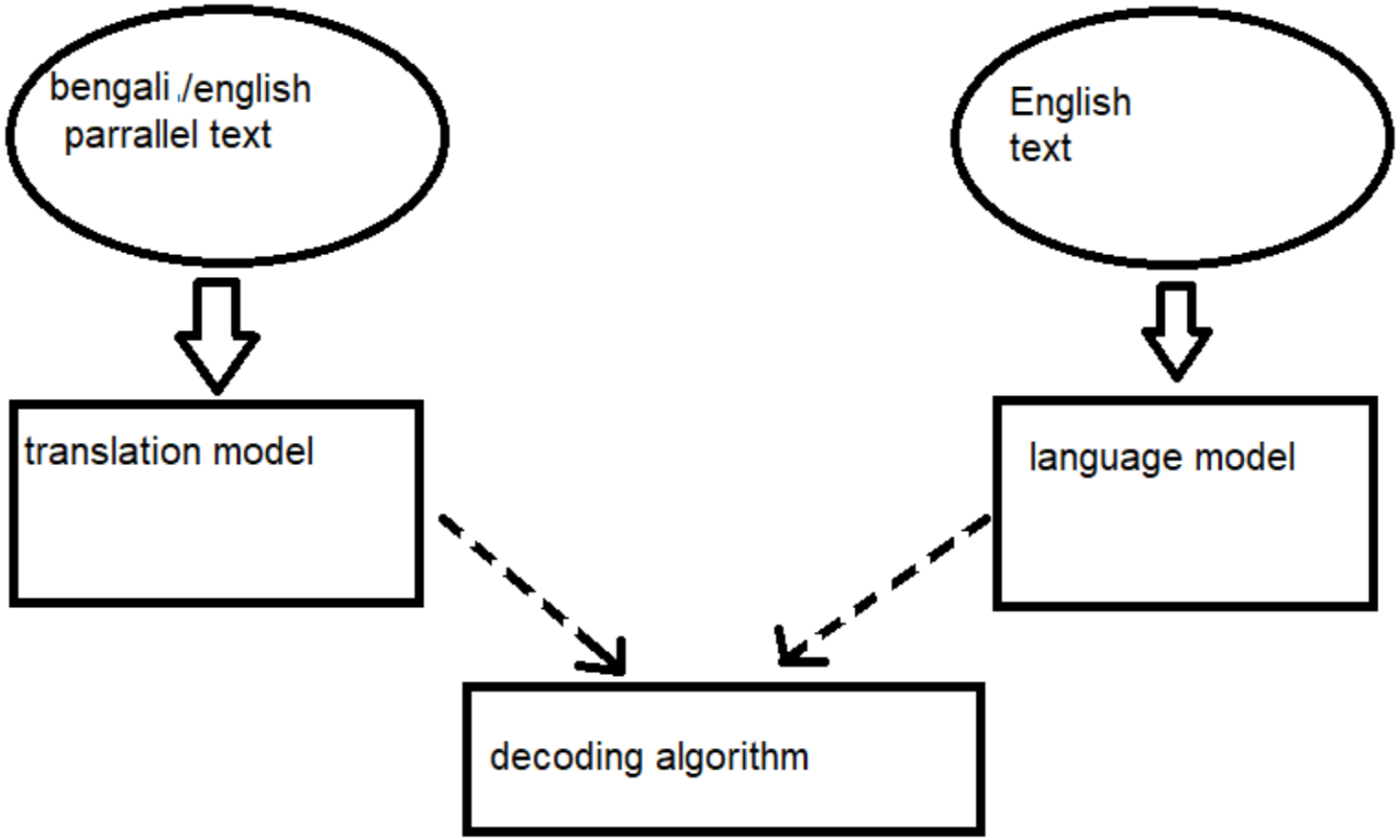
MT Contd…..

- It is conditional probability distribution P(Y|X) where Y is the target language representation given X as source language.
- We can say given a source language X what is the maximum probability of target language Y i.e. maximizing the probability of P(Y|X) in statistical machine translation.
- This is data driven approach and hence requires corpus of source and target language examples

MT System contd.: Drawback Phrase based SMT

- The main drawback of Phrase based statistical machine translation is it does not consider the context.
- Only language model of statistical machine translation considers context.
- There are mainly two important objectives of any machine translation, adequacy and fluency.
- Adequacy refers to the measure of portion of translated text with respect to input text.
- Fluency is the measure of grammatically correctness of translated text with respect to the input text.

MT Contd….SMT Architecture.



MT Contd……: Hybrid MT model

- **Hybrid Translation Model:** Hybrid model combines rule based and statistical machine translation engine to generate high performance and improved translation quality.

MT Contd……: NMT

- Neural Machine Translation(NMT) Neural Machine translation uses the concept of Artificial Neural Network. Recently NMT exploits the concept of deep learning to translate source to target language.
- NMT directly targets the conditional probability P(x|y) to model it for translating a source sentence $x_{1,\ldots\ldots}x_n$ to target sentence $y_{1,\ldots\ldots}y_{n..}$ then finally decoder decodes one target word at a time from the entire source language. It is done by decomposing the conditional probability

MT Contd….: NMT

- The NMT do not require any domain specific knowledge and linguistic features for source and target language.
- Also NMT systems jointly train its encoder and decoder model unlike phrase based SMT systems where different components are trained and tuned separately to get optimal performance.
- Unlike SMT model where special hand crafted features are linearly combined for translation, NMT model automatically learns from distributed representation of data in non linear transfer function under trained network .

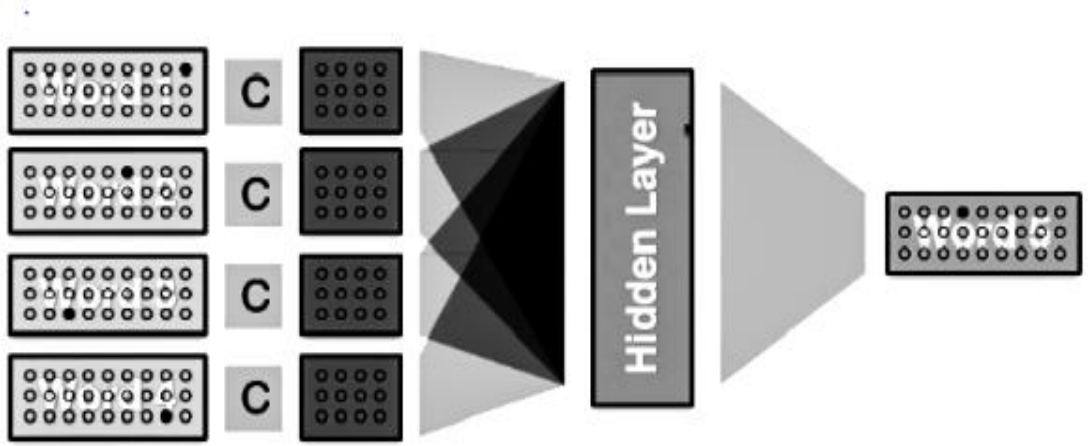
MT Contd…: NMT

- The first thing that is Neural language model with feed forward network is how to represent words in neural network nodes.
- All nodes in the neural networks carry real numbered values.
- Even we can’t simply use token ids because neural network will assume that for token 12133 is similar to token 12134 but practically these numbers are completely arbitrary in nature.
- The solution : each word with high dimensional vector and representing one dimension per word in the vocabulary and the value 1 for the dimension that matches with the word, and assigning 0 for the rest. This type of vector representation is called one hot vector.
- For example:
- Cat=(0,0,0,0,0,1,0,0,0……..)

MT Contd….: NMT

- These are very large vectors for the huge vocabulary size.
- To pool evidence, there is another layer exists in between input layer and hidden layer.
- In this layer, in low dimensional space each context word is represented. This is known as word embedding.

MT Contd……NMT

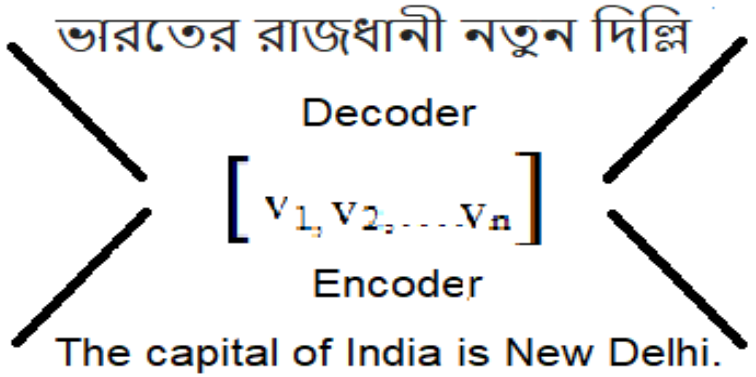


MT Contd…..

- The above diagram(Fig.) represents the feed forward neural language model.
- Context words W1, W2, W3,W4 are represented as one hot vector , then it is projected into continuous space with the help of word embedding using weight matrix C for all words.
- The predicted word w5 is computed with the help of hidden layer and the output predicted word is also one hot vector

MT Contd……: NMT

- **Encoder Decoder based Neural Machine Translation:**



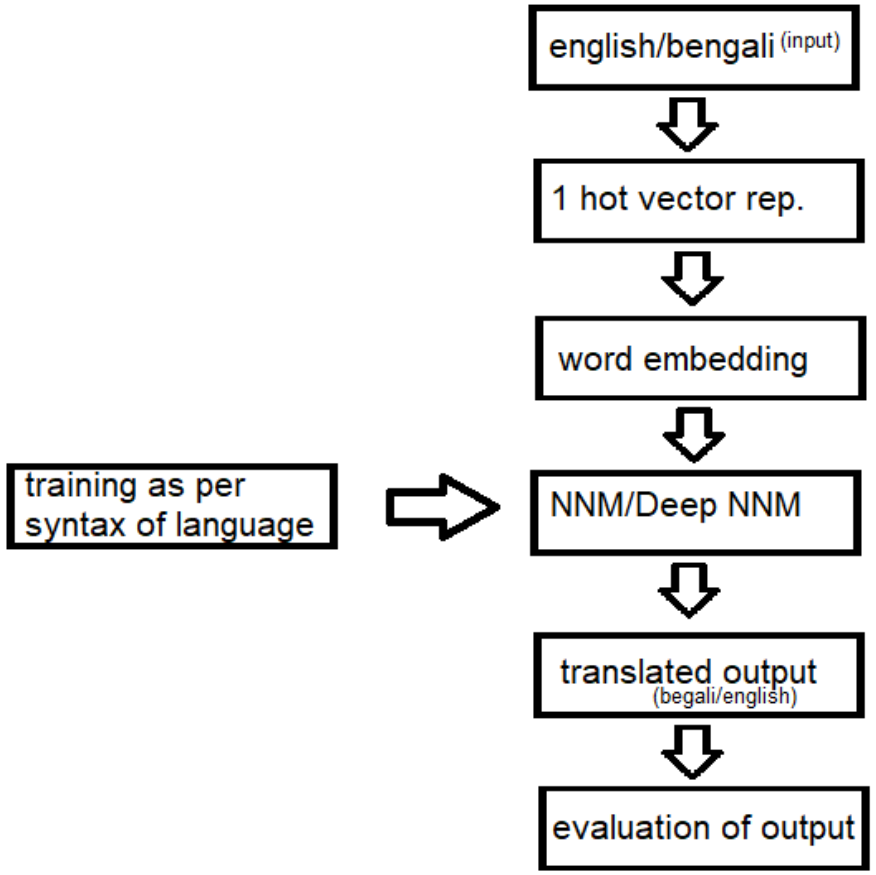
NMT contd…..Methodology

- Starting with the baseline model of Statistical Machine Translation for which Moses will be used. For NMT we will use Sockeye which is mainly based on MAXNET, Sequence to sequence NMT using Pytorch, Open NMT etc.
- First we will require bilingual parallel corpora. In our case we need English and Hindi corpus. We may require to preprocess our corpus if needed. Next we have to train the Neural network model with the help of linguistic syntax of the corresponding language.
- For example, in English the syntactic structure of the sentence is subject verb object(SVO). But in Bangla it is Subject object verb(SOV).
- Typical steps to be followed are as under:
- Step 1: Train the SMT/Neural network model using selected bilingual corpora.

MT Contd……: NMT Methodology

- Step 2: Validate the different MT models with validation data set and tuning different hyperparameters.
- Step 3: Test the model with test data.
- Step 4 : Evaluate the result with automatic evaluators like BLEU, METEOR,NIST etc. and finally with the help of human evaluator best model will be proposed in terms of adequacy and fluency.

NMT Contd……: Methodology



Machine Learning in Health Care

- Identifying diseases and diagnosis
- Drug discovery and manufacturing
- Medical Imaging diagnosis
- Clinical trial and research
- Outbreak prediction etc. etc.

Identifying diseases and diagnosis

- IBM Watson genomics is a best instance of how to rapidly diagnose genome based cancer cell replication.
- BERG(AI powered biotech company) the biopharmaceutical organization uses AI in fields such as oncology to produce innovative medicine. It uses the power of AI ,coupled with patient biology and thereby accelerate clinical identification and subsequently treating the disease. Hence a faster discovery and development of treatment.
- Medical imaging diagnosis: The ground breaking technology called computer vision was responsible for both machine learning and deep learning.
- Outbreak prediction: ProMED mail a web based monitoring system that tracks the occurrence of infectious disease and offers outbreak alerts in live time.

Role of ML in drug discovery

- The drug development and production pipelines are large, complicated and rely on many considerations.
- ML methods have plentiful , high quality data resources that can improve discovery and judgment making.
- Examples : Target verification, clinical trial analysis of pathology data etc.
- Various classifiers like SVM, Decision tree, ensemble methods, naïve bayse classifiers,KNN,ANN etc. help in this context

Medical Image Analysis

- AI imaging tools can screen chest X-rays for signs of tuberculosis, often achieving a level of accuracy comparable to humans. This capability could be deployed through an app available to providers in low-resource areas , reducing the need for a trained radiologist on site.
- Electronic Health Record(HER) providers are now using AI to create more intuitive interfaces and automate some of the routine processes that consume so much of users time. Electronic Health Record data can help to identify infection patterns and highlight patients at risk before they begin to show symptoms.
- Voice recognition and dictation are helping to improve the clinical documentation process , but NLP tools are helpful.

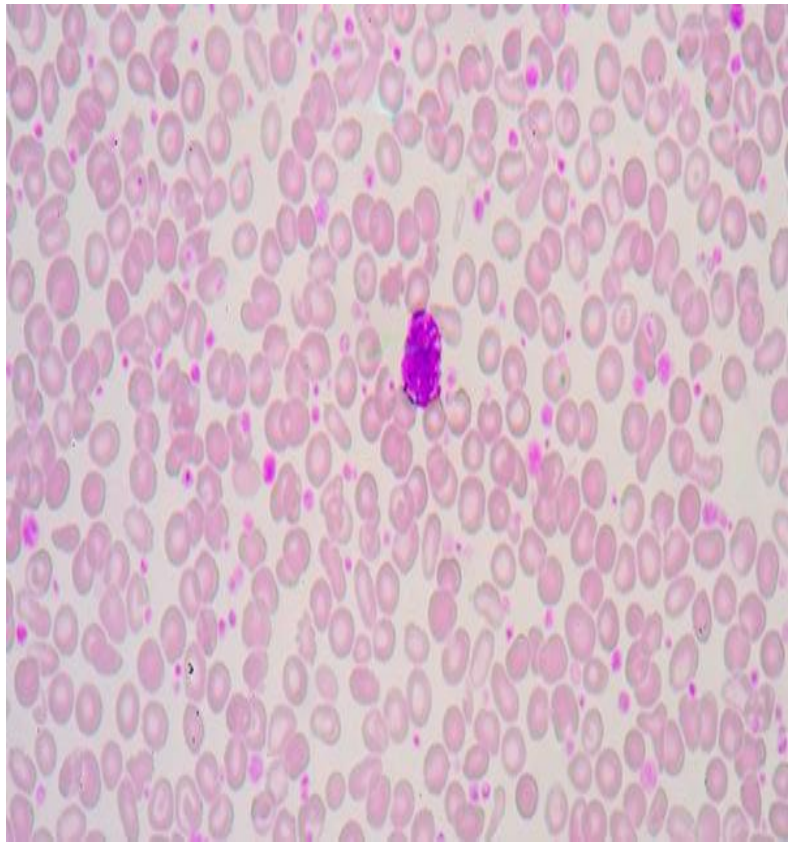
Segmentation and Detection

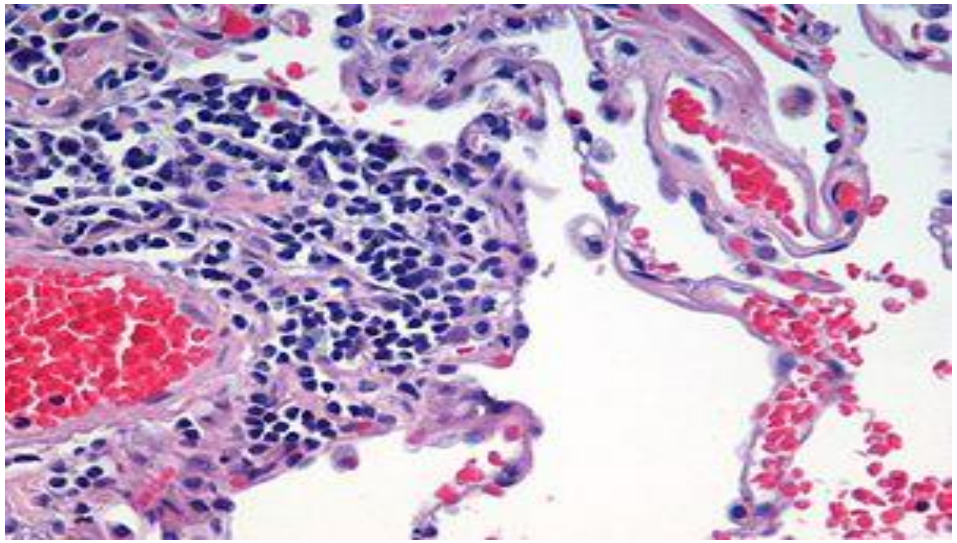
- Digital Pathology (DP) is the process by which histology (study of tissues/cells)slides are digitized to produce high resolution images.
- These digitized slides afford the possibility of applying image analysis techniques to DP for applications in detection, segmentation and classification.
- A segmentation task is defined as the requirement of delineating an accurate boundary for histologic primitives(i.e., nuclei etc…)so that precise morphological features can be extracted.
- Detection tasks are different from segmentation tasks in that the goal is to typically to simply identify the center of the primitive of interest and not explicitly extract the primitive contour or boundary.

Segmentation and Detection

- Identifying individual tissue based primitives and trying to identify primitive specific feature/patches to make predictions regarding tissue class, an alternative strategy is to directly learn the set of features representatives of the tissue class via DL.
- The DL classifiers could be trained to self discover the disease pattern within each class.

Medical Image Analysis

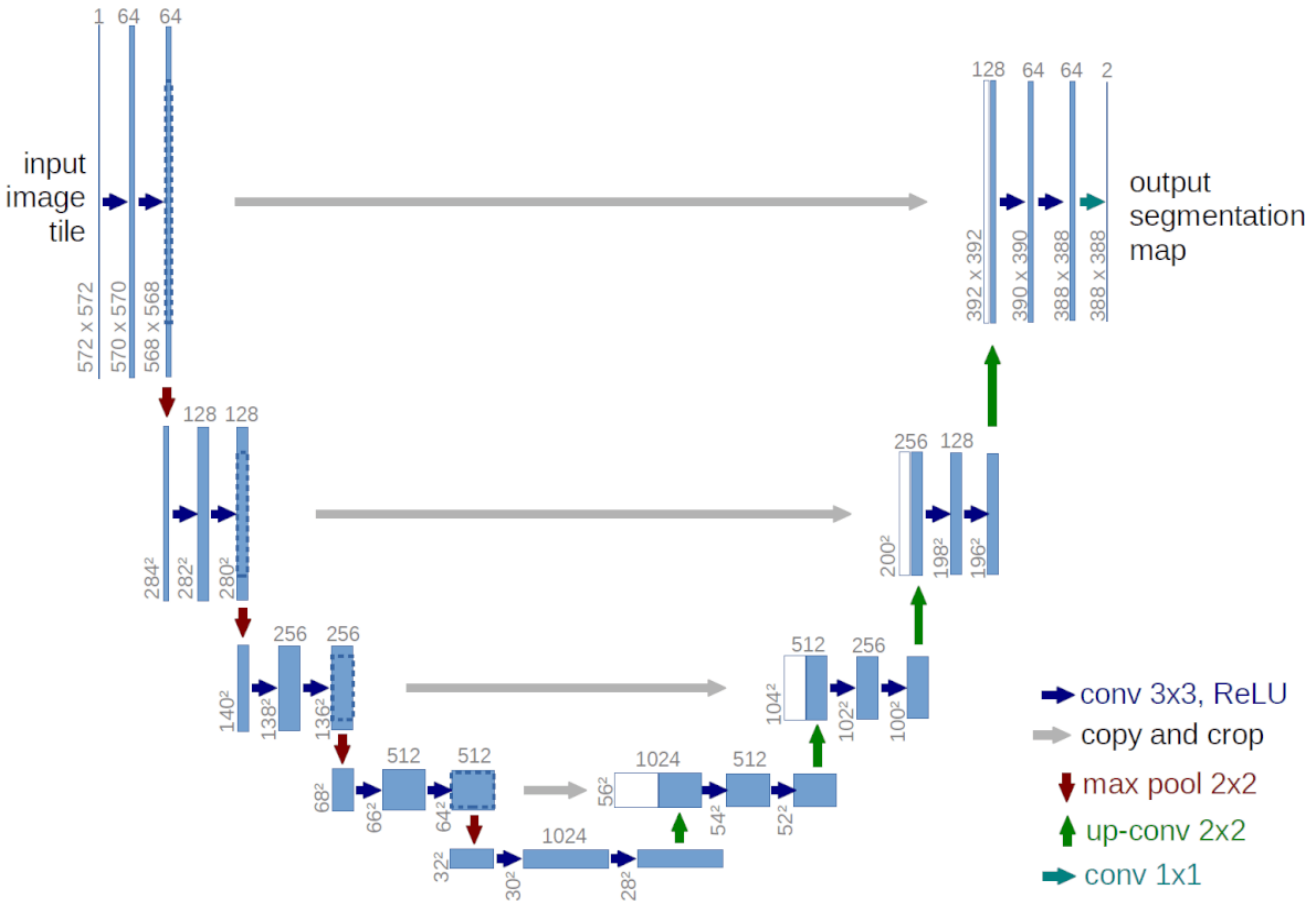




U-Net Architecture:

- U-Net is a convolutional NN architecture.
- U-Net is used for Image segmentation.
- It uses multiple channel feature maps.
- It is mainly used for image segmentation and also to process biomedical images.
- In biomedical images it not only distinguish whether there is a disease but also to find out area of abnormality.
- U-net network is strong enough and could do good prediction with small data sets.

U-Net Architecture

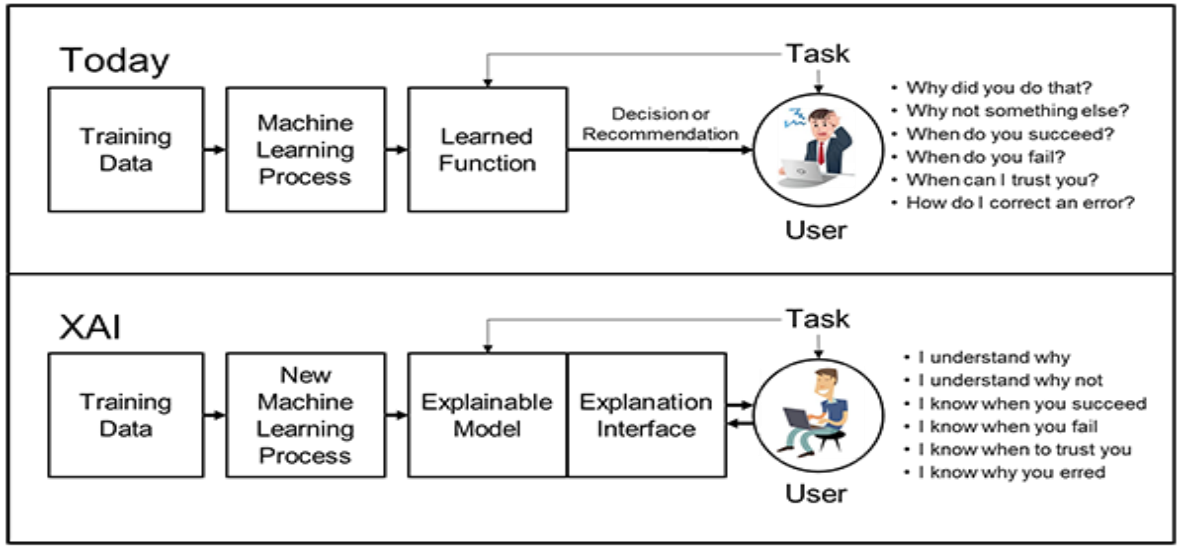


Framework for AI/ML based Computational Histopathology(Microscopic examination of tissue) Analysis .

- Data Acquisition and Preparation : Annotation by pathologists –region of interest , cell marking

- AI/ML processing : Trained with the patches and features are classified with ML based classifiers such as SVM, random forest etc. or DL based classifiers.
- Results and Visualizations of output : Validation and testing of individual models (Performance measures- Accuracy, precision,recall,F1-Score etc.)

Explainable AI



Explainable AI

- We are entering a new age of AI applications …….
- ML is the core technology
- ML models are opaque, non intuitive, and difficult for people to understand

End users can probably get the answers of the following:

Why did you do that ?

Why not something else??

When did you succeed???

When do you fail????

When can I trust you ????? And many more…….???????????

ML in Insurance and Retail Industries.

- Data: Identifying unmet needs…
- Decisioning: Using customer behavior to trigger effective campaigns.
- Predictive model for insurance underwriting :
- Insurance underwriting is the process of evaluating a company's risk in insuring a home,car,or an individual's health or life.
- It determines whether it would be profitable for an insurance company to take a chance on providing insurance coverage to an individual or business.
- Life Insurance underwriting is all about risk.
- Risk level of a person is directly proportional to the premium he/she will pay.
- If person is less risky than the average to insure , he will pay less and if his risk level is higher than the average he will pay more.

Insurance Underwriting

- Underwriter uses software and finally approves or rejects the risk of issuing a policy…..

Loan underwriting rules:

- During underwriting, the lender gauges the credit worthiness of the borrower and assesses whether the applicant meets the loan eligibility criteria or not.
- Credit Score:
- Income:
- Current debt and liabilities :
- Value of the Collateral(The term collateral refers to an asset that a lender accepts as security for a loan):
- Property appraisal :

Regression and Correlation, Clustering and Classification algorithms are widely used.

- Both bivariate regression and correlation can be used to evaluate the strength of a relationship between two variables.
- Regression is generally used to predict future values based on the past values by fitting a set of points to a curve.
- Correlation used to examine the degree to which the values for two variables behave similarly.

Association Rule /Market Basket/ Buying Pattern/Retail Store

- The purchasing of one product when another product is purchased represents an association rule.
- Association rules are frequently used by retail stores to assist in marketing, advertisement, floor placement and inventory control.
- Association rules are directly applicable to retail business.
- Association rules are used to show the relationship between data items.

Sample data to illustrate Association Rules:

• Transactions	Items
T1	Bread, Jelly, PeanutButter
T2	Bread, PeanutButter
T3	Bread, milk, PeanutButter
T4	Beer, Bread
T5	Beer, Milk

Retail Market/ Market Basket Model

- The support (s) for an association rule X → Y is the percentage of transactions in the database that contain X U Y.
- Confidence/Strength (α) for an association rule X → Y is the ratio of the number of transactions that contain XUY to the number of transactions that contain X
- For Example : bread → Peanutbutter , support=60%, Confidence=75%
- Peanut butter → Bread , Support=60%, Confidence=100%
- Beer → Bread , Support =20%, confidence=50%
- Peanutbutter → Jelly , Support=20%, confidence=33.3 %

Association Rule Mining : Basic Algorithms

- Apriori Algorithms : Apriori algorithm is the most well known association rule algorithm and is used in most commercial products…
- ANY SUBSET OF A LARGE ITEM SET MUST BE LARGE.
- Sampling Algorithms :
- Partitioning Algorithms :
- Parallel and Distribution Algorithms…..

Rule based fraud detection vs ML based fraud detection

- Rule-Based fraud detection:

❑Catching obvious fraudulent scenarios

❑Requires much manual work to enumerate all possible detection rules

❑Multiple verification steps may become threat to user experience.

❑Long term processing

ML based fraud detection

- ❑Finding hidden and implicit correlations in data
- ❑Possible fraud scenario detection happens automatically
- ❑Reduction in the number of verification measures
- ❑Real time processing

Applications of ML and artificial intelligence in transportation

- The ML travel programs were implemented in a variety of forms.

Among them for instance:

- Use machine learning techniques in strategic choice-making and management.
- AI technologies are also discussed with a view to improving public transport.
- Connected and autonomous cars are the next promising AI technology in transport, which tries to increase efficiency by limiting the amount of road disaster.

Applications of machine learning in transport

- AI technique can be viewed as an intelligent remedy for these complicated systems that traditional methods cannot be used to manage.
- A lot of AI methods are also used in transportation such as ANNs.
- Few applications of machine learning in transport are:

❑Incident detection.

❑ Predictive models.

Incident detection

- There have been many efforts to classify the date, location, and extent of an accident and assist traffic management in reducing congestion.
- Using statistical techniques such as the , event detection algorithms were first introduced.
- Algorithms have been developed for neural network approaches.
- An element-oriented model of a recurring time-lag system (TLRN-Time lag recurrent system) neural network has been developed.

Application of AI in aviation and public transportation

- Aviation : To control the flight trip more efficiently, AI has been acknowledged.
- The safety of the aircraft was measured by testing the engine on board utilizing the Probabilistic Neural Network (PNN)
- Bus Service : ANNs is easy to use by forecasting bus arrival times to reduce the waiting time for passengers. Bus Arrival, departure etc.
- An Automatic Vehicle Location (AVL) system aimed at improving the performance of public transport activities, retaining operational controls and enhancing the overall quality of public transport services

Full Syllabus-End Term Examination:

- **Unit 1: A Brief Introduction to Machine Learning**
- Paradigms, Knowledge Representation, Data Acquisition, Data Pre-Processing, Feature Extraction and Processing Feature Ranking and Selection, Feature Reduction, Model Learning, Evaluation and Deployment
- MT Translation System and its evaluation: Statistical to Neural , BLEU
- **Unit 2: Machine Learning in Banking and Securities**
- **Unit 3: Machine Learning in Communication, Media and Entertainment**
- **Unit 4: Machine Learning in Healthcare and Life Sciences**
- **Unit 5: Machine Learning in Education**
- **Unit 5: Machine Learning in Insurance**
- **Unit 6:Machine Learning in Transportation and Logistics**