Reinforcement Learning & NLP

Reinforcement Learning: Introduction, Passive Reinforcement Learning, Active Reinforcement Learning, Generalization in Reinforcement Learning, Policy search, applications of RL

Natural Language Processing: Language models, Text classification, Information Retrieval, Information Extraction.

---

1. Reinforcement Learning:

Reinforcement learning is a feedback based learning technique in which an agent learns to behave in an environment by performing the actions and seeing the results of actions.

For each good action, the agent gets positive feedback (or reward), and for each bad action, the agent gets negative feedback (or penalty).

In RL, the agent learns automatically using feedbacks without any labeled data, the agent learns by its experience only.

"Reinforcement learning is a type of machine learning method where an intelligent agent (computer program) interacts with the environment and learns to act within that".

The primary goal of an agent in RL is to improve the performance by getting the maximum +ve rewards.

It is a core part of AI, all AI agents works on the concept of RL. We do not need to ~~perform~~

pre-program the agent.

→ Take action
→ change state / Remain in the same state
→ Get feedback.

By performing above three actions, he learns and explores the environment.

↓ agent

Reinforcement Learning Terminology:-

Agent:- An entity that can percepts & explores the environment by performing actions on it

Environment:- A situation in which an agent is present & surrounded by.

In RL, we assume the stochastic environment, which means it is random in nature.

Action:- Actions are the moves taken by an agent within the environment.

State:- State is a situation returned by the environment after performing action by an agent.

Reward:- A feedback returned to the agent from the environment to evaluate the action of the agent.

Policy:- It is a strategy applied by the agent for the next action based on the current state.

Value:- It is opposite to the short-term reward and expected long-term returned with the discount factor.

Q-value:- It is similar to the value, but it takes one additional parameter as a current action.

In RL, the agent takes the next action and changes states according to the feedback of the previous action.

---

2. Approaches to Implement RL:-

There are 3 approaches to implement RL in ML.

→ Value based
→ Policy based
→ Model based

The value based approach is about to find the optimal value function, which is the maximum value at a state under any policy.

The policy based approach is to find the optimal policy for the max future rewards without using the value function. It has 2 types of policy.

→ Deterministic
→ Stochastic.

Deterministic: The same action is produced by the policy at any state.

stochastic: In this policy, probability determin... the produced action.

In the model based approach, a virtual mode... is created for the environment, and the age... explores the environment to learn it. There i... no particular solution / algorithm for this approach.

Elements of RL:-

There are 4 main elements of RL..

→ Policy
→ Reward signal
→ Value Function
→ Model of the environment.

Policy:-

Policy is a way how an agent behaves at a given time. It maps the perceived states of the environment to the actions taken on those states. Policy can define the behaviour of an agent.

→ Deterministic policy = $\pi(s)$
→ Stochastic policy = $\pi(a|s)$

Reward Signal:-

The goal of RL is defined by the reward signal. At each state, the environment sends ~~a~~ signal to the agent, and this signal is called a reward signal.

these rewards are given according to the good and bad actions taken by the agent.

Value function:-

The value function gives information about how good the situation and action are and how much reward an agent can expect.

A value function specifies the good state and action for the future.

The value function depends on the reward, if without reward, there could be no value.

Model:- model, which minics the behavior of the environment.

With the help of the model, one can make interfaces about how the environment will behave. Such as, if a state and an action are given, then a model can predict the next state and reward.

The approaches for solving the RL problems with the help of model are termed as the model-based approach. An approach without using a model is called model free approach.

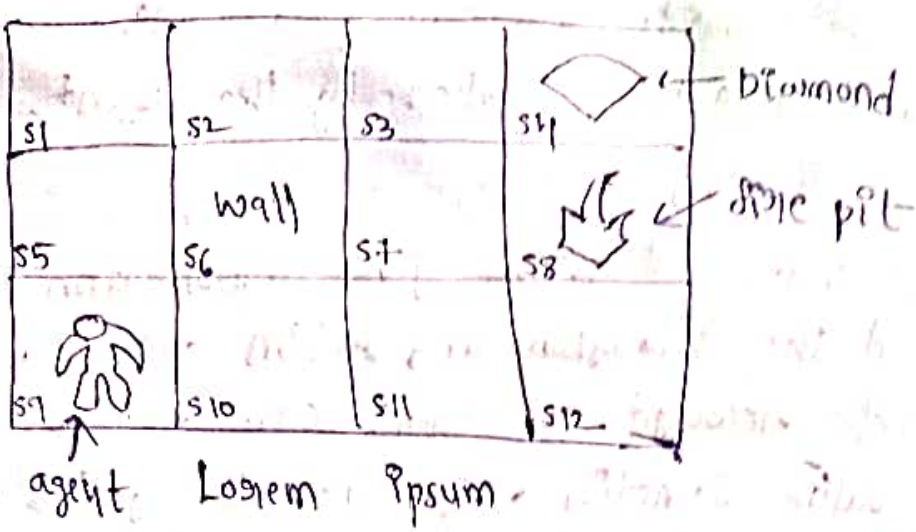
---

3 Working process of RL:-

To understand the working process of RL, we consider 2 things.

Environment & agent.

Agent is an intelligent program such as AI robot.



Lorem Ipsum.

The agent can not cross the S_6 block, as it is a solid wall. If the agent reaches the S_4 block, then get the +1 reward, if it reaches the fire pit, then it gets the -1 reward point.

It can take four actions: move up, move down, move left, move right.

The agent will try to remember the preceding steps that it has taken to reach the final step. To memorize the steps, it assigns 1 value to each previous step.

S1 V=1	S2 V=1	S3 V=1	S4 Diamond
S5 V=1	S6 Wall	S7	S8 Fire
S9 Agent V=1	S10	S11	S12

If the agent in S1 block, it get confused, because, whether he should go ~~up~~ ~~down~~ down or right, S_2 and S_5 blocks have same value V=1.

To solve this problem, we will use Bellman equation, which is the main concept behind RL.

The Bellman Equation :-

The Bellman equation was introduced by the mathematician Richard Ernest Bellman in the year 1953, hence it is called Bellman equation. It is associated with dynamic programming.

The Bellman equation is given by,-

$$V(s) = \max \left[R(s,a) + \gamma V(s') \right]$$

where,

$V(s)$ = value calculated at particular point.

$R(s,a)$ = reward at particular state (s) by performing an action(a).

γ = Discount factor

$V(s')$ = The value at the previous state

We are taking the maximum value, because the agent tries to find the optimal solution.

at S_3 block

$$V(S_3) = \max \left[R(s,a) + \gamma (V(s')) \right]$$

$R(s,a) = 1$, reward is diamond.

$V(s') = 0$, no further state to move

$\gamma = 0.9$

$$V(S_3) = \max \left[1 + 0.9 (0) (0) \right] = \max [1+0] = 1$$

Ily at S_2 block

$$V(S_2) = \max \left[0 + 0.9 (1) \right] = 0.9 = V(S_7)$$

$R(s,a) = 0$, no reward

$\gamma = 0.9$, $V(s') = 1$, There is a move.

$V(s12) = \max [0 + 0.9(0.9)] = 0.81 = V(s11)$

$V(s5) = \max [0 + 0.9(0.81)] = 0.73 = V(s10) = V($

$V(s9) = \max [0 + 0.9(0.73)] = 0.66$

0.81 →	0.9 →	1 →	Diamond
↑ 0.73	Wall	0.9	Fire pit
↑ 0.66 Agent	0.73	0.81	0.73

4. Types of RL:-

There are mainly 2 types of RL.

(i) Positive RL

(ii) Negative RL

Positive RL:-

The positive RL means adding something to increase the strength of the behavior and the behavior of the agent.

Too much positive RL may lead to an overload of states.

Negative RL:-

The negative RL is opposite to the positive RL

It can be more effective than the positive RL depending on situation and behavior.

We can represent the agent state by using the Markov state that contains all the required information from the history.

The Markov state follows the ~~agent state~~ Markov property, which says that the future is independent of the past and can only be defined with the present.

Markov Decision Process:-

MDP is used to formalize the RL problems. In MDP, the agent constantly interacts with the environment and perform actions, at each action, the environment responds and generates a new state.
MDP is used to describe the environment for the RL. MDP uses Markov property.

Markov Property:-

"If the agent is present in the current state S1, performs an action a1 and move to the state S2, then the state transition from S1 to S2 only depends on the current state and future action and do not depend on the past actions, rewards or states.

(or

Markov property, the current transition does not depend on any past action or state.
Ex: chess game

Markov Process:-

It is a memoryless process with a sequence of random states $S_1, S_2, \ldots, S_t$ that uses the Markov property. Markov process also known as Markov chain.

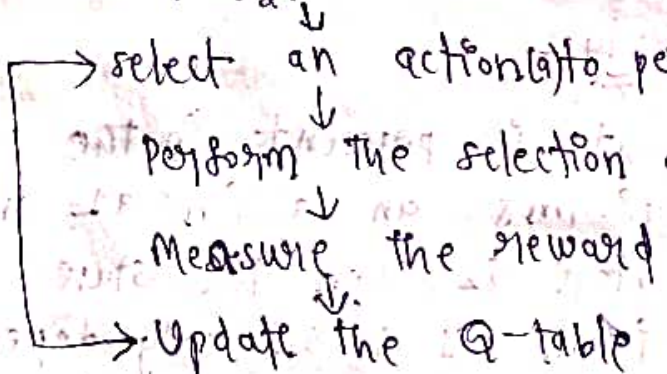
5. RL Algorithms :-

RL algorithms are used in AI applications and gaming applications. The mainly used algorithms are

→ Q learning
→ SARSA - State Action Reward State Action.
→ DQN - Deep Q Neural Network.

Q Learning :-

Q learning is an Off policy RL algorithm, which is used for the temporal difference learning.



SARSA :-

It is an on-policy temporal difference learning method.

In SARSA, new action and reward are selected using the same policy.

The SARSA uses the quintuple (s, a, r, s, a).

where, s = original state
a = original action
r = reward
s = new state
a = new action.

DQN :-

DQN is a Q-learning uses Neural Networks. For big state space environment, it will be a challenging and complex task to define and update Q-table.

To solve this issue, we can use a DQN algorithm. Instead of using Q table, neural network approximates the Q values for each action and state.

Q-learning Explanation:-

Q-learning is a popular model-free RL algorithm based on the Bellman equation.

The goal of the agent in Q-learning is to maximize the value of Q.

In Q-learning, Q- stands for Quality of an action.

Q-table :-

The Q-table in matrix is created while performing the Q-learning. The table follows the state and action pair.

The RL agents interact with the environment, explore it, take action, and get rewarded.

Supervised learning algorithms learn from the labeled dataset and on the basis of the training, predict the output.

6. RL Applications :-

The applications of RL are

- → Game playing
- → Chemistry.
- → Finance sector
- → Manufacturing
- → Business
- → Robotics
- → Control.

(i) Robotics : RL is used in Robot navigation, Robo-soccer, walking etc.

(ii) Control : RL used for adaptive control such as Factory process, admission control in telecommunication etc.

(iii) Game playing : RL is used in Game playing such as tic-tac-toe, chess etc.

(iv) Chemistry : RL is used for optimizing the chemical reactions.

(v) Business : RL is used for business strategy planning.

(vi) Manufacturing : In automobile manufacturing companies, the robots are using deep RL to pick goodes and put them in containers.

(vii) Finance sector : RL is used in finance sector for evaluating trading strategy

1. NLP :-

NLP stands for Natural Language Processing, which is a part of Computer Science, Human language and AI.

It is the technology that is used by machines to understand, analyse, manipulate, and interpret human language.

It helps developers, to organize knowledge for performing task such as translation, automatic summarization, Named Entity Recognition (NER), Speech Recognition, Relationship Extraction, and topic segmentation.

History of NLP :-

The Natural Languages Processing started in the year 1940s.

In 1948, the first recognisable NLP application was introduced. In Birkbeck college London.

In 1960s-80s flavored with AI. such as

→ Augmented Transition Networks (ATN)

ATN is a finite state machine that is capable of recognizing regular languages.

→ Case Grammer was developed by Linguist Charles J. Fillmore in the year 1968.

It is used for expressing the relationship between nouns and vers by using the preposition.

SHRDLU is a program written by Terry Winograd in 1968-70. It helps users to communicate with the computer and moving objects.

LUNAR is the classic example of NL DB interface system that is used ATNs and Woods's Procedural semantics.

NLP consists of various applications like speech recognition, machine translation and machine text reading etc.

Eg:- Amazon ALEXA.

Advantages of NLP:-

→ NLP helps users to ask any question and get direct response within seconds.
→ NLP offers exact answers to the question.
→ NLP helps computers to communicate with humans in their languages.
→ It is very time efficient.
→ Most companies uses NLP to improve the efficiency of documentation process, accuracy documentation etc.

Disadvantages of NLP:-

→ NLP may not show context.
→ NLP is unpredictable.
→ NLP may require more keystrokes.

---

2. Components of NLP:-

There are 2 components of NLP.

→ NLU → NLG.

NLU :-

NLU – Natural Language Understanding.

NLU helps to machine to understand and analyse the human language.

NLU mainly used in Business applications to understand the customers problem in both spoken and written language.

NLG :-

NLG – Natural Language Generation.

It acts as an translator that converts the computerized data into natural language representation.

The NLG is difficult than NLG.

Applications of NLP :-

The applications of NLP are

(i) Question Answering.

Building systems that automatically answer the questions asked by humans in NL.

ex:- Alexa, Google assistant, Cortana, Siri etc.

(ii) Spam Detection.

It is used to detect unwanted emails getting to a users inbox.

(iii) Sentiment Analysis.

It is also known as opinion mining.

It is used on the web to analyse the attitude, behaviours, and emotional state of the sender.

(iv) Machine Translation
It is used to translate text or speech from one language to another language.
Ex:- Google translator.

(v) Spelling Correction
Microsoft provides word processor like MS-word and pp for the spelling correction.

(vi) Speech Recognition.
It is used for converting spoken words into text.
Ex: Voice biometric.

(vii) Chatbot.
It is an important application of NLP.
It is used by many companies to provide the customer's chat services.

(viii) Information Extraction.
It is used for extracting structured information from unstructured or semi-structured machine-readable documents.

(ix) NLU
It converts large set of text into more formal representations.

---

3. NLP Pipeline:-

There are the following steps to build NLP pipeline.

(1) Sentence segmentation.
It breaks the paragraph into separate sentences.

(ii) Word Tokenization.

It is used to break the sentence into seperate words (or tokens.

(iii) Stemming.

It is used to normalize words into its base form (or root form.

Eg:- celebrates, celebrated, celebrating.

rootword - celebrate.

(iv) Lemmatization

It is quite similar to stemming.

It is used to group different inflected forms of the word, called lemma.

It produces the root word, which has a meaning.

(v) Identifying stop words.

Is, in, was, the, and, a, an etc, these are the stop words in english.

Stopwords might be filtered out before doing any statistical analysis.

Note: When you are building a rock band search engine, then you do not ignore the word "The".

(vi) Dependency Parsing.

It is used to find, how all the words in a sentence are related to each other.

(vii) POS tags.

POS stands for Parts Of Speech.

A word has one or more parts of speech based on the content. Eg:- Google something in the net.

(viii) NER

Named Entity Recognition is the process of detecting the named entity such as person name, movie name, organization name etc.

Ex:- Steve Jobs, Bill Gates, Microsoft, SKD etc.

(ix) Chunking.

It is used to collect the individual piece of information and grouping them into bigger pieces of sentences.

---

4. Phases of NLP:-

There are the following 5 phases of NLP.

(i) Lexical and Morphological Analysis.

It scans the source code as a stream of characters and converts it into meaningful lexemes.

It divides the whole text into paragraphs, sentences and words.

(ii) Syntactic Analysis (Parsing)

It is used to check grammar, word arrangements and shows the relationship among the words.

(iii) Semantic Analysis.

It is concerned with the meaning representation. It mainly focus on the literal meaning of words, phrases and sentences.

(iv) Discourse Integration.
It depends upon the sentences that proceeds it and also invokes the meaning of the sentences.

(v) Pragmatic Analysis.
It helps you to discover the intended effect by applying a set of rules that characterize cooperative dialogues.

NLP is difficult because ambiguity in the language.

(i) Lexical Ambiguity:
A single word has a two (or more meanings.
Ex:- Manya is looking for a match.

(ii) Syntactic Ambiguity.
A sentence has 2 (or more possible meanings.
Ex:- I saw the girl with the binocular.

(iii) Referential Ambiguity.
It exists when you are referring to something using the pronoun.

NLP APIs:-

NLP APIs allow developers to integrate human to machine communications.

→ IBM Watson API.
→ Chatbot API
→ Speech to text API
→ Sentiment Analysis API

→ Translation API by SYSTRAN
→ Text analysis API by AYLIEN
→ Cloud NLP API
→ Google Cloud Natural Language API

NLP Libraries:-

→ Scikit - learn
→ Natural Language Toolkit (NLTK)
→ Pattern
→ TextBlob
→ Quepy
→ Spacy
→ Gensim

---

5. Text Classification:-

Text classification is the process of categorizing the text into a group of words.
By using NLP, text classification can automatically analyze text. NLP is used for sentiment analysis, topic detection, and language detection.

There are 3 types of text classification approaches

→ Rule based System
→ Machine System
→ Hybrid System.

In the rule based approach, texts are seperated into an organized group using a set of handcraft linguistic rules.
Those handicraft linguistic rules contain users to define a list of words that are categorized by groups.

Ex:- Ronaldo, Virat ⇒ Sportsmans
Obama, Modi ⇒ Politicians

Machine based classifier learns to make a classification based on past observation from the data sets. User data is prelabled as train and test data.

It collects the classification strategy from the previous inputs and learns automatically.

Machine based classifier usage a bag of words for feature extension.

In a bag of words, a vector represents the frequency of words in a predefined dictionary of a word list.

We can perform NLP using Naive Bayes, SVM. and Deep Learning algorithms.

Text → Feature Extractor → Features → ML algo^m → classifier Model.

Hybrid approach usage combines a rule-based and machine based approach.

It uses the rule based system to create a tag and use ML algo^{m}s to train the system to create a rule.

The machine-based rule list is compared with the rule-based rule list.

If something does not match on the tags, humans improve the list manually.

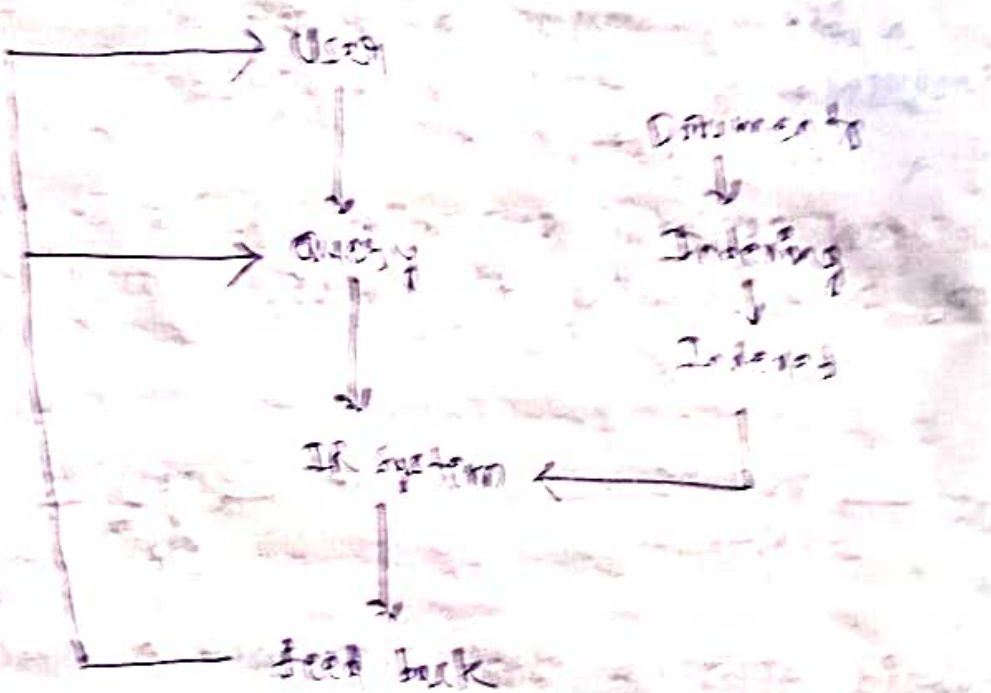
It is the best method to implement text classification.

Information Retrieval :-

Information Retrieval (IR) may be defined as a software program that deals with the organization, retrieval and evaluation of information from document repositories particularly text and information.

The system assists users in finding the information they require but it does not explicitly return the answers of the questions.

A modern IR system will maintain only references to documents.



Classical Problem :-

The main goal of IR research is to develop a model for retrieving information from the repositories of documents.

A classical problem named ad-hoc retrieval problem related to the IR system.

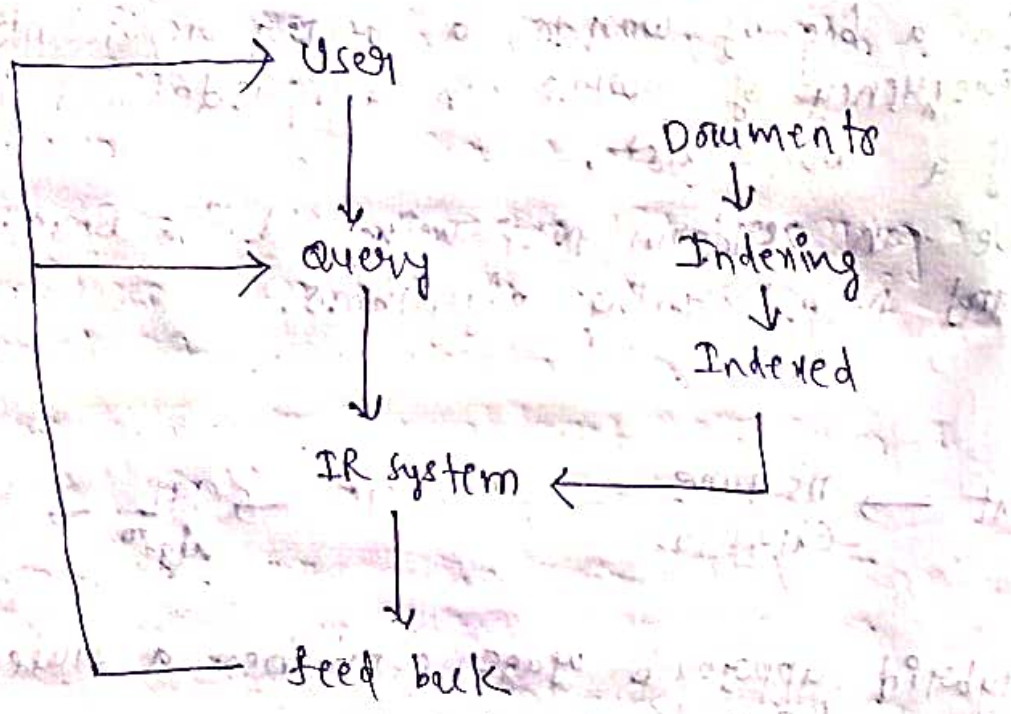
In ad-hoc retrieval, the user enters a query in natural language that describes the

6. Information Retrieval :-

Information Retrieval (IR) may be defined as a software program that deals with the organization, storage, retrieval and evaluation of information from document repositories particularly textual information.

The system assists users in finding the information they require but it does not explicitly return the answers of the questions.

A perfect IR system will retrieve only relevant documents.



Classical Problem :-

The main goal of IR research is to develop a model for retrieving information from the repositories of documents.

A classical problem named ad-hoc retrieval problem related to the IR system.

In ad-hoc retrieval, the user must enter a query in natural language that describes the

required information. Then the IR system will return the required documents related to the desired information.

IR Model :-

A model of information retrieval predicts and explains what a user will find in relevance to the given query.

IR model is basically a pattern that defines the following (or consists the following

→ A model of documents

→ A model of queries

→ A matching function that compares queries to documents.

Types of IR Models :-

The IR model can be classified into the following three models.

Classical IR model :-

It is simplest and easy to implement IR model. This model is based on mathematical knowledge that was easily recognized and understood as well.

→ Boolean → vector → probabilistic

are the 3 classical IR models.

Non-classical IR model :-

It is completely opposite to classical IR model.

Ex :- Information logic model
Situation Theory Model
Interaction models.

Alternate IR model :-

It is the enhancement of classical IR model making use of some specific techniques from other fields.

Ex: cluster model, Fuzzy model
Latent Semantic Indexing (LSI) model.

Design features of IR system :-

(i) Inverted Index : It is the primary data structure of most of the IR systems.

(ii) Stop Word Eliminations.

Stop words are high frequency words. They have less semantic weights.

Ex:- a, an, the, in, of, for, at etc.

As per zipf's law, a stop list covering a few dozen words reduces the size of inverted index by almost half.

If we eliminate the alphabet "A" from "Vitamin A" then it would have no significance.

(iii) Stemming : It is the simplified form of morphological analysis.

It is the heuristic process of extracting the base form of words by chopping off the ends of words.

Ex: laughing - laughs - laughed = laugh.

The Boolean Model :-

It is the oldest IR model. The model is based on the set theory and the boolean algebra where documents are set of terms and queries are boolean expressions on terms.

The boolean model can be defined as

D - Documents

Q - Queries ($\wedge$ $\vee$ $\sim$)

F - Framework for D & Q.

R - Ranking.

Ex: ((text $\vee$ information) $\wedge$ retrieval $\wedge$ $\sim$ theory)

$\wedge$ → Intersection (AND) *

$\vee$ → Union (OR) +

Advantages :-

→ The simplest model, which is based on sets.
→ Easy to understand and implement.
→ It only retrieves exact matches.

Disadvantages :-

→ No ranking for retrieved documents.
→ The query language is expressive, but it is complicated too.
→ Boolean operator has much more influence than a critical word.

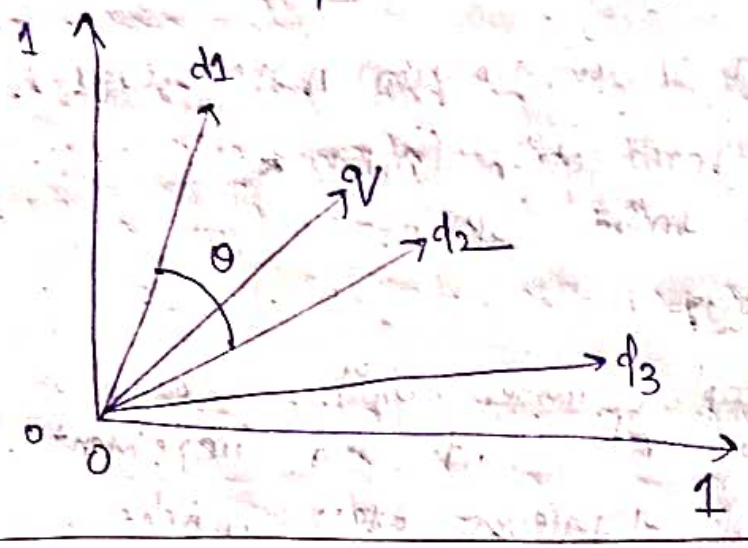
Vector Space Model :-

Due to the above disadvantages of the boolean model, Gerard Salton and his colleagues suggested a model, which is based on Luhn's similarity criterion.

→ The index representations (documents) and the queries are considered as vectors embedded in a high dimensional Euclidean space.

→ The similarity measure of a document vector to a query vector is usually the cosine of the angle b/w them.

The query and documents are represented by a 2D vector space.



7. Information Extraction :-

Information Extraction is the process of parsing through unstructured data and extracting essential information into more editable and structured data formats.

Information extraction from text data can be achieved by leveraging Deep Learning and NLP techniques like NER - Named Entity Recognition.

How Does IE Work :-

The understand the mechanism of IE using NLP algorithms we should understand the kin of data we are working on.

This will help us to sort out the information we want to extract from the unstructured data.

Ex: while working on medical reports, it should identify and extract patient names, drug information, and other general reports.

Below are some of the common techniques to build models.

Tokenization:

Computers usually won't understand the language we speak to communicate.

The process of breaking down language into tokens is called tokenization.

Ex: NLP Information Extraction is fun.

One-word (Unigram Token): NLP, Information, Extraction, is, fun.

Two-word phrase (Bigram Tokens): NLP Information, Information Extraction, Extraction is, is fun.

Three-word sentences (Trigram Tokens):

NLP Information Extraction, Information E is, Extraction is fun.

POS Tagging:

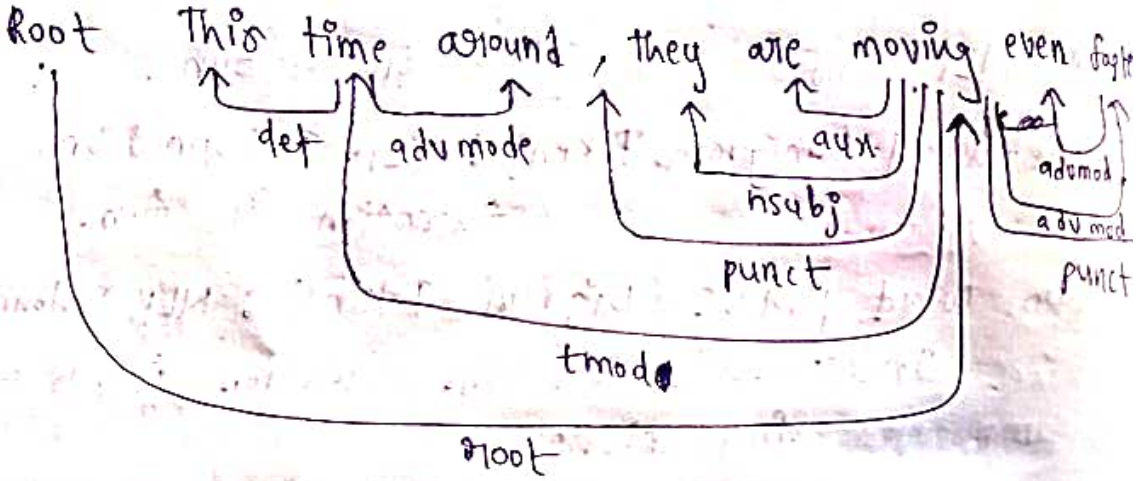
POS tagging is very crucial for information extraction from text. It will help us understand the context of the text data.

We refer the text data from the documents as unstructured data. Hence, we use pos tagging technique.

In pos tagging, all the tokens in the text data get categorized into different word categories such as nouns, verbs, adjectives, prepositions, determiners etc.

Dependency Graphs:-

This help us for finding the relationships blw neighbouring words using directed graphs. This relation will provide details about the dependency type (eg. subject, object etc).



NER with Spacy:-

Spacy is an open-source NLP library for advanced NLP in Python and Cython. Information Extraction with Spacy NER models are widely leveraged.