

Naive Bayes is a probabilistic algorithm which is based on probability theory and Bayes' Theorem to predict the tag of a text such as news or customer review. It helps to calculate the probability of each tag for the given text and return the tag with the highest probability. Bayes' Theorem is used to predict the probability of a feature based on prior knowledge of conditions that might be related to that feature. The choice of area in NLP using Naïve Bayes Classifiers could be in usual tasks such as segmentation and translation but it is also explored in unusual areas like segmentation for infant learning and identifying documents for opinions and facts. Anggraeni et al. (2019) [61] used ML and AI to create a question-and-answer system for retrieving information about hearing loss. They developed I-Chat Bot which understands the user input and provides an appropriate response and produces a model which can be used in the search for information about required hearing impairments. The problem with naïve bayes is that we may end up with zero probabilities when we meet words in the test data for a certain class that are not present in the training data.

(b) Hidden Markov Model (HMM)

An HMM is a system where a shifting takes place between several states, generating feasible output symbols with each switch. The sets of viable states and unique symbols may be large, but finite and known. We can describe the outputs, but the system's internals are hidden. Few of the problems could be solved by Inference A certain sequence of output symbols, compute the probabilities of one or more candidate states with sequences. *Patterns matching* the state-switch sequence are most likely to have generated a particular output-symbol sequence. *Training* the output-symbol chain data, reckon the state-switch/output probabilities that fit this data best.

Hidden Markov Models are extensively used for speech recognition, where the output sequence is matched to the sequence of individual phonemes. HMM is not restricted to this application; it has several others such as bioinformatics problems, for example, multiple sequence alignment [128]. Sonnhammer mentioned that Pfam holds multiple alignments and hidden Markov model-based profiles (HMM-profiles) of entire protein domains. The cue of domain boundaries, family members and alignment are done semi-automatically found on expert knowledge, sequence similarity, other protein family databases and the capability of HMM-profiles to correctly identify and align the members. HMM may be used for a variety of NLP applications, including word prediction, sentence production, quality assurance, and intrusion detection systems [133].

(c) Neural Network

Earlier machine learning techniques such as Naïve Bayes, HMM etc. were majorly used for NLP but by the end of 2010, neural networks transformed and enhanced NLP tasks by learning multilevel features. Major use of neural networks in NLP is observed for word embedding where words are represented in the form of vectors. These vectors can be used to recognize similar words by observing their closeness in this vector space, other uses of neural networks are observed in information retrieval, text summarization, text classification, machine translation, sentiment analysis and speech recognition. Initially focus was on feedforward [49] and CNN (convolutional neural network) architecture [69] but later researchers adopted recurrent neural networks to capture the context of a word with respect to surrounding words of a sentence. LSTM (Long Short-Term Memory), a variant of RNN, is used in various tasks such