Module-2 (part-2)

Text Modelling: Bayesian Networks, Hidden Markovian Models, Markov random Fields, Conditional Random Fields

Bayesian Networks

- Bayesian Networks, also known as Bayesian Belief Networks or Bayes Nets, are probabilistic graphical models that represent uncertain relationships between variables using a directed acyclic graph (DAG). They are named after the Reverend Thomas Bayes, a mathematician who developed Bayes' theorem, which forms the foundation of Bayesian Networks.
- Bayesian Networks model the conditional dependencies between variables in a probabilistic domain. They represent a joint probability distribution over a set of random variables by decomposing it into a set of conditional probability distributions.
- The structure of a Bayesian Network is defined by a directed acyclic graph (DAG), where nodes represent random variables, and directed edges represent conditional dependencies between variables.
- Each node in the graph corresponds to a random variable, and its state is influenced by its parent nodes, which are the variables it depends on.

Components:

- Nodes: Represent random variables or observable events in the domain of interest.
- Edges: Represent probabilistic dependencies between variables. An edge from node A to node B indicates that B is conditionally dependent on A.
- Conditional Probability Tables (CPTs): Store the conditional probability distributions of each node given its parent nodes. These tables specify the probabilities of each possible state of the node given the states of its parents.

Applications

- Medical Diagnosis: Bayesian Networks are used for diagnosing diseases based on symptoms and medical test results. They can combine evidence from multiple sources to provide a probabilistic assessment of the patient's condition.
- Risk Assessment: In finance and insurance, Bayesian Networks are employed for risk assessment and decision-making. They model the dependencies between risk factors and help in estimating the likelihood of adverse events.
- Natural Language Processing: Bayesian Networks are applied in tasks such as language modeling, part-of-speech tagging, and sentiment analysis. They capture dependencies between words or linguistic features and aid in probabilistic inference.
- Fault Diagnosis: In engineering and manufacturing, Bayesian Networks are utilized for fault diagnosis and troubleshooting. They model the relationships between components and symptoms to identify the root causes of failures.

Hidden Markov Models (HMMs

- Hidden Markov Models (HMMs) are statistical models used to model sequential data where the underlying system is assumed to be a Markov process with unobservable (hidden) states. HMMs have been widely applied in various fields, including speech recognition, natural language processing, bioinformatics, and finance. HMMs are based on the concept of a Markov process, where a system transitions between a finite set of states over discrete time steps.
  1. In an HMM, the states of the system are unobservable (hidden), but each state generates an observable output (emission) with a certain probability.
  2. The model assumes the Markov property, meaning that the probability of transitioning to the next state depends only on the current state and not on the previous history of states.

- Hidden States: Represent unobservable states of the system that evolve over time according to the Markov property.
- Observations (Emissions): Represent observable outputs generated by each hidden state at each time step.
- Transition Probabilities: Specify the probabilities of transitioning from one hidden state to another. These probabilities are represented by a transition matrix.
- Emission Probabilities: Specify the probabilities of emitting each observation given the current hidden state. These probabilities are represented by an emission matrix.

Applications:

- **Speech Recognition:** HMMs are widely used in speech recognition systems to model the temporal dynamics of speech signals and recognize spoken words or phonemes.
- **Natural Language Processing:** In NLP, HMMs are applied to tasks such as part-of-speech tagging, named entity recognition, and machine translation, where sequential data modeling is required.
- **Bioinformatics:** HMMs are used for analyzing biological sequences such as DNA, RNA, and protein sequences. They are employed in tasks like gene finding, sequence alignment, and protein structure prediction.
- **Finance:** HMMs are utilized in finance for modeling time series data such as stock prices, interest rates, and economic indicators. They are applied in areas like risk management, portfolio optimization, and algorithmic trading.

- **Speech Recognition:** HMMs are widely used in speech recognition systems to model the temporal dynamics of speech signals and recognize spoken words or phonemes.
- **Natural Language Processing:** In NLP, HMMs are applied to tasks such as part-of-speech tagging, named entity recognition, and machine translation, where sequential data modeling is required.
- **Bioinformatics:** HMMs are used for analyzing biological sequences such as DNA, RNA, and protein sequences. They are employed in tasks like gene finding, sequence alignment, and protein structure prediction.
- **Finance:** HMMs are utilized in finance for modeling time series data such as stock prices, interest rates, and economic indicators. They are applied in areas like risk management, portfolio optimization, and algorithmic trading.

Markov Random Fields (MRFs)

- Markov Random Fields (MRFs) are probabilistic graphical models used to represent complex dependencies among variables in a given domain. MRFs are characterized by an undirected graph structure where nodes represent variables, and edges represent dependencies or interactions between variables.
- MRFs model the joint probability distribution of a set of random variables by defining a set of local interactions between neighboring variables in an undirected graph.
- They are based on the Markov property, which states that the probability distribution of each variable depends only on its neighbors in the graph.

Cont..

- **Nodes (Vertices):** Represent random variables or elements of interest in the domain. Each node corresponds to a variable that we want to model or make inferences about.
- **Edges:** Connect pairs of nodes and represent the dependencies or interactions between variables. The absence of an edge between two nodes indicates conditional independence given all other variables in the graph.
- **Factors (Potentials):** Functions defined over subsets of variables in the graph. They capture the relationships between variables and determine the strength of their interactions.

- **Nodes (Vertices):** Represent random variables or elements of interest in the domain. Each node corresponds to a variable that we want to model or make inferences about.
- **Edges:** Connect pairs of nodes and represent the dependencies or interactions between variables. The absence of an edge between two nodes indicates conditional independence given all other variables in the graph.
- **Factors (Potentials):** Functions defined over subsets of variables in the graph. They capture the relationships between variables and determine the strength of their interactions.

Applications:

- **Image Processing:** MRFs are widely used in computer vision and image processing tasks such as image denoising, image segmentation, and image restoration. They model the spatial dependencies between pixels in images and improve the accuracy of these tasks.
- **Social Network Analysis:** MRFs are applied in social network analysis to model interactions between individuals in a network. They capture dependencies between nodes (individuals) and can be used for tasks such as community detection, link prediction, and influence analysis.
- **Natural Language Processing:** MRFs are utilized in NLP for tasks such as text summarization, machine translation, and syntactic parsing. They model dependencies between words or linguistic features and enable structured prediction.
- **Remote Sensing:** In remote sensing applications, MRFs are used for image classification, land cover mapping, and change detection. They model dependencies between pixels in remote sensing images and improve the accuracy of these tasks.

Conditional Random Fields (CRFs)

- Conditional Random Fields (CRFs) are a type of discriminative probabilistic graphical model used for structured prediction tasks, particularly in sequential data modeling. CRFs are an extension of Hidden Markov Models (HMMs) and Markov Random Fields (MRFs), designed to address some of their limitations.
- CRFs model the conditional probability of a set of output variables (labels) given a set of input variables (features).
- Unlike generative models like HMMs, which model the joint distribution of input and output variables, CRFs directly model the conditional distribution of output variables given input variables.
- CRFs are discriminative models, meaning they focus on learning the decision boundary between different output labels rather than modeling the entire joint distribution.

- **Input Features:** Represent observed data or input variables that provide information for predicting the output labels.
- **Output Labels:** Represent the variables we want to predict or infer. These labels are typically structured and sequential in nature.
- **Feature Functions:** Define the relationship between input features and output labels. They capture the compatibility between input features and potential label assignments.
- **Parameters:** CRFs have parameters associated with feature functions, which are learned from training data using techniques such as maximum likelihood estimation or gradient descent.

Sample Question

- Explain Bayesian Networks in details.
- Elaborate Hidden Markovian Models and their advantage.
- Explain Markov random Fields.
- Short note on Conditional Random Fields