International Journal of Academic and Applied Research (IJAAR)
ISSN: 2643-9603
Vol. 8 Issue 2 February - 2024, Pages: 52-65

Knowledge Representation Using Statistical and Probabilistic Reasoning

Bardi Ifeanyi Akanna [1] and Irechukwu Obiageli Jacquelin[2]

1Information and Communication Technology Department
Alvan Ikoku University of Education, Owerri
bardi.ifeanyi@alvanikoku.edu.ng
2Information and Communication Technology Department
Alvan Ikoku University of Education, Owerri
Irechukwu.obiageli@alvanikoku.edu.ng

Abstract: *Developing models and structures to efficiently capture and organize information for reasoning and problem-solving, a crucial aspect known as knowledge representation, holds significant importance in the realms of artificial intelligence and cognitive science. In recent times, the utilization of statistical and probabilistic reasoning has emerged as powerful technique for effective knowledge representation. This article employs these methods to delve into diverse facets of knowledge representation, encompassing fundamental principles, methodologies, applications, and challenges. As real-world problems grow in complexity, knowledge representation has undergone a transformation to encompass more nuanced and robust decision-making processes. This study explores the paradigm shift towards knowledge representation utilizing statistical and probabilistic reasoning, presenting a more adaptable approach by integrating uncertainty into knowledge models. Unlike conventional methods like symbolic logic, which grapple with challenges in handling incomplete information and uncertainty, the newer approaches prove more adept. Theoretical foundations and practical applications of statistical and probabilistic reasoning, including but not limited to Bayesian networks, Markov networks, and influence diagrams, are thoroughly examined across various domains such as healthcare, finance, and natural language processing. These models facilitate probabilistic knowledge representation and enhance decision-making processes, rendering them invaluable for sound reasoning under conditions of uncertainty.*

Keywords: Knowledge Representation, Statistical and Probabilistic Reasoning

Introduction

In the expansive realm of artificial intelligence and machine learning, effective knowledge representation stands as a cornerstone for the success of intelligent systems. Knowledge representation involves capturing, organizing, and structuring information to enable machines to reason, make decisions, and solve intricate problems. According to Branchman and Bector (2014), knowledge representation is a dynamic area of research in artificial intelligence, often involving the technical efforts of knowledge engineers in acquiring domain knowledge for knowledge-based systems.

An influential approach to knowledge representation involves statistical and probabilistic reasoning techniques, as highlighted by Zhang,Q (2015). This approach utilizes probability to signify the level of uncertainty in information, effectively combining probability theory with logic to handle uncertainty arising from incomplete or uncertain data. It proves particularly valuable when faced with uncertainties about the truth of statements, enabling the assignment of a numerical measure of uncertainty to each proposition. This numerical data is then combined using uniform syntactic principles.

The representation of knowledge is a vital aspect of artificial intelligence, allowing machines to make decisions based on available information. Probabilistic reasoning, employing probability to indicate uncertainty, becomes crucial in managing uncertain information. Statistical and probabilistic reasoning methods are essential for handling complex, uncertain, and imprecise real-world data, making AI systems more adaptable.

Statistical reasoning, involving the interpretation and summarization of statistical information (Benjamin, 2019), becomes a crucial part of these methods. Statistical and probabilistic reasoning empower us to model and manage uncertainty, providing valuable applications in healthcare, finance, natural language processing, robotics, and autonomous systems. These methods enable the quantification of uncertainty, informed decision-making with incomplete or noisy data, and the development of intelligent and robust systems.

This exploration of knowledge representation using statistical and probabilistic reasoning delves into key concepts, methodologies, and applications in the field. Various formalisms and techniques, such as Bayesian networks, Markov models, and statistical learning methods, are examined for representing and reasoning with uncertain knowledge. The practical implications of these approaches in decision support, pattern recognition, and prediction across diverse domains are also explored.

The journey through knowledge representation using statistical and probabilistic reasoning uncovers the synergy between data-driven methodologies and intelligent systems. It reveals the advantages of enhancing our understanding of intricate information and simulating real-world situations. Our goal is to unveil the intricate framework of knowledge representation and its groundbreaking influence on the future of artificial intelligence. Additionally, we discuss the obstacles and future prospects in the field, utilizing statistical and probabilistic reasoning.

Statistical and probabilistic reasoning's increasing significance in knowledge representation provides a flexible and potent approach to modeling uncertainty and making informed decisions. As knowledge representation evolves, these techniques will play a critical role in addressing the challenges presented by today's data-rich and uncertain world.

The Key Concepts: Knowledge Representation and Reasoning

Comprehending the principles of knowledge representation and reasoning is vital for the optimal functioning of artificial intelligence (AI). Knowledge in AI denotes the information possessed by a system, sourced from data collection, human input, or existing knowledge bases. Efficient management and utilization of this knowledge are imperative for machines to make well-informed decisions and execute tasks. Knowledge can be expressed through various means, including symbolic representations, statistical models, and semantic networks, with the accurate and efficient representation being pivotal for the success of AI systems.

Representation in AI pertains to how information is organized and stored to enable machines to access and comprehend it. A proficient representation enables AI systems to model and reason about the world, drawing conclusions from the available information. Different methodologies employ diverse representation methods, such as rule-based systems, neural networks, or ontologies, depending on the specific problem. The choice of representation significantly influences an AI system's performance and its ability to adapt to new information and scenarios.

Reasoning involves the process of making inferences or drawing conclusions based on the knowledge and representations available to an AI system. In AI, reasoning techniques can be deductive, inductive, abductive, probabilistic, or statistical, depending on the nature of the problem and the available information. Effective reasoning is crucial for AI systems to make decisions, solve problems, and generate meaningful outputs.

Concept of Knowledge

A deep understanding of statistical and probabilistic reasoning holds significance across various domains, spanning from data science and machine learning to decision-making within uncertain environments. Statistical reasoning involves the capacity to scrutinize data, discern patterns, and draw informed inferences about populations or phenomena based on observed samples. This skill set facilitates the derivation of meaningful conclusions and the quantification of uncertainties in assessments. On the other hand, probabilistic reasoning addresses uncertainty, employing probability theory to quantify uncertainties. This skill enables the modeling and prediction of outcomes in situations characterized by randomness and unpredictability.

Proficiency in statistical and probabilistic reasoning is a valuable asset for individuals and AI systems alike, enhancing the ability to make well-informed choices, mitigate risks, and leverage the potential of data for extracting valuable insights. In our data-driven world, this skill set emerges as crucial for navigating complexities and making informed decisions.

Concept of Representation

In the realm of artificial intelligence and cognitive science, representation pertains to the structuring and encoding of information for processing by machines or cognitive systems. This fundamental concept plays a crucial role, as the manner in which data or knowledge is represented significantly affects a system's ability to comprehend, reason, and manipulate information effectively. Various types of representations are employed in AI, and the selection of a specific representation can deeply influence the system's performance and capabilities.

There Are Various Forms of knowledge Representation In AI, Including:

Symbolic Representation: This method employs symbols or discrete entities to express knowledge, often associated with rule-based systems and logic-based reasoning. It involves defining explicit rules and relationships symbolically, making it well-suited for domains with clear, structured knowledge.

Connectionist Representation: In contrast to symbolic representation, connectionist or neural network-based representation utilizes distributed and continuous representations. These systems rely on interconnected nodes (neurons) to collectively capture and process information, excelling at tasks such as pattern recognition and learning from data.

Semantic Representation: This form of representation concentrates on capturing the meaning of data, typically utilizing structured ontologies or semantic networks. Semantic representations aim to encode knowledge in a manner that reflects relationships and hierarchies between concepts.

Statistical Representation: Statistical models and representations play a significant role in machine learning. They capture patterns and statistical dependencies in data, making them suitable for tasks like classification, regression, and probabilistic reasoning.

Geometric Representation: In certain AI applications, especially in computer vision and robotics, geometric representations are employed to model spatial relationships and objects in the environment. This can involve representations like point clouds or 3D models.

Concept of Reasoning

It involves formal manipulation of symbols expressing a set of beliefs to reflect new ones. We leverage the notion that symbols are more accessible than propositions here. Statistical and probabilistic reasoning help robots make judgements and draw conclusions under uncertainty in artificial intelligence and data science. Statistics are used to analyse data, find trends, and make predictions. AI systems use hypothesis testing, regression analysis, and clustering to get insights from data. However, probabilistic reasoning uses probability theory to express and manipulate uncertainty. It is especially useful when judgements must be based on probabilities when the result is unknown. Bayesian networks and Markov models are used to model and update beliefs, evaluate risks, and make logical, probabilistic judgements. Statistical and probabilistic reasoning help AI systems make better decisions in real-world situations with inadequate, noisy, or uncertain data.

The Significance of Knowledge Representation

Effective knowledge representation underpins AI systems. It lets robots interpret and manipulate data like humans. AI applications including natural language processing, computer vision, autonomous systems, recommendation engines, and more need this capacity. It involves acquiring, organising, and structuring data for computational reasoning and problem-solving. Machine reasoning, learning, and complicated problem solving depend on knowledge representation. Ontologies, semantic networks, and logic-based formalisms have been used to represent and organise knowledge.

knowledge is essential for AI systems for several reasons:

1. Problem Solving: Knowledge helps AI systems grasp and solve complicated problems. AI can use relevant knowledge to make judgements and solve problems using a broad variety of inputs.
2) Contextual Understanding: Knowledge lets AI systems grasp task or conversation context. Context is essential for meaningful replies and natural language comprehension.
3. Generalisation: Knowledge lets AI systems generalise from examples to anticipate or decide in new scenarios. Generalisation is necessary for AI to be flexible.
4. Learning: AI systems use machine learning to find patterns and correlations in data. AI systems may learn from knowledge, identifying important data patterns and connections.
5. Reasoning: Knowledge aids AI reasoning. It helps people solve problems and make decisions by using reasoning, drawing conclusions, and inferring.
6. Safety and Ethics: Understanding ethical principles, regulations, and social standards is essential for ethical AI systems. It prevents AI from producing harmful or biassed results.
7. Human-AI Interactions: Knowledge helps AI systems answer user questions. It helps them answer questions, make suggestions, and have meaningful interactions.
8. Domain-specific Expertise: AI systems require domain-specific expertise to execute jobs in medical, finance, and engineering. This expertise may include processes, rules, or industry-specific terminology.
9. Problem Domain Understanding: AI systems like autonomous cars, robotics, and natural language processing must comprehend their problem domain. Safe and successful operation requires domain knowledge about the environment, objects, and entities.
10. Adaptation and Evolution: Knowledge helps AI systems develop. AI systems may enhance their performance by updating their knowledge base with fresh information.

AI system capabilities and effectiveness depend on knowledge. It lays the groundwork for comprehending, reasoning, learning, and interacting with people and the environment, making AI more versatile.

Traditional knowledge representation

Rule-based systems and ontologies struggle with unclear, partial, or ambiguous data. Statistical and probabilistic reasoning help here. Machines can reason under uncertainty and make educated real-world judgements using these strategies.

Probabilistic reasoning

Probabilistic reasoning addresses uncertainty and probability. It entails drawing conclusions from information and the probability of distinct outcomes concerning unclear occurrences or circumstances. This reasoning evaluates the probability or likelihood of certain occurrences.

Some probabilistic reasoning ideas are:

1. Probability Theory: Probabilistic reasoning is based on probability theory. Probability theory gives guidelines for assessing uncertainty and formulating probabilistic predictions.

2. Bayesian Inference: Bayesian inference updates probability based on fresh evidence using Bayes' theorem. It includes updating beliefs or probability with fresh evidence.

3. Uncertainty Modelling: Incomplete information, unpredictability, and randomness create uncertainty in many real-world situations. Probabilistic reasoning lets decision-makers explicitly express and manage uncertainty.

4. Decision Making under Uncertainty: Probabilistic reasoning is employed in uncertain decision-making. It lets decision-makers evaluate outcomes and make educated decisions based on success or failure possibilities.

5. Machine Learning and AI: Probabilistic reasoning underpins many algorithms. Bayesian networks, probabilistic graphical models, and other probabilistic approaches describe uncertainty and forecast in diverse applications.

Statistics: Probabilistic reasoning is similar to statistical reasoning. Sample data is used in statistics to assess probability and make population inferences.

Probabilistic reasoning is used in weather forecasting, medical diagnosis, financial modelling, and autonomous systems that make uncertain judgements.
Probabilistic thinking helps people handle uncertainty and make reasonable judgements in unpredictable circumstances.

Statistical Reasoning

Using statistical tools and concepts, statistical reasoning helps people make informed judgements and draw meaningful inferences from data. Data collection, organisation, and analysis reveal patterns, trends, and correlations that may be used to anticipate or support hypotheses. Statistics is essential in research, economics, healthcare, and social sciences since it quantifies uncertainty and evaluates results. It explains the complicated and unpredictable world by offering a systematic and logical framework for evaluating information, making choices, and solving real-life issues.
Statistical reasoning represents and reasons with knowledge using statistical models and methods. This method works well with huge datasets and noisy data.

Examples of Statistical Models and Techniques That Are Used in Various Fields to Represent and Analyze Data

Linear Regression

Linear Regression: Linear regression involves fitting a linear equation to observed data to depict the relationship between a dependent variable and one or more independent variables. An illustrative example is predicting property values based on factors such as square footage, bedrooms, and location.

Logistic Regression: Logistic regression addresses binary classification problems with binary outcomes (0,1). An example scenario is predicting binary outcomes like subscription churn based on usage trends.

Example: Predicting subscription churn based on use trends.
Choice Trees
We employ decision trees for categorization and regression. They segment a dataset by characteristics.
Example: Using credit score, income, and work history to predict loan default.
Random Forest
Random forests use numerous decision trees to increase prediction accuracy and avoid overfitting.
Example: Predicting e-commerce client preferences using browsing history, demographics, and purchase history.

K-Means Grouping
K-means clustering groups related data points unsupervisedly.
Using shopping behaviour to segment clients for targeted marketing.
Principal Component Analysis
PCA reduces high-dimensional data while keeping variance.
Example: Dataset dimensionality reduction for visualisation or feature selection.
Support Vector Machines
A supervised learning approach for classification and regression is SVM. It seeks the optimum hyperplane to classify data points.
Classifying emails as spam or not based on content and properties.
Time-series analysis
Time series analysis models and forecasts data across time.
Forecasting stock prices, weather, or demand in supply chain management.

Neural Networks
Image identification, natural language processing, and voice recognition employ neural networks, including deep learning models like CNNs and RNNs.
Example: Computer vision system picture classification or language model text generation.
Bayesian Networks
Bayesian networks display and reason about data uncertainty using probabilistic graphics.
Example: Modelling a patient's illness likelihood based on symptoms and medical history.
Statistics descriptions
Descriptive statistics use mean, median, and standard deviation. These statistics help explore and comprehend data distributions by revealing central tendency, dispersion, and shape.
Statistical Inference
Based on a sample, inferential statistics determine a population. From a subset of data, hypothesis testing and confidence intervals enable researchers to generate population-wide predictions.
Statistic Learning
Data patterns and correlations are extracted using regression, clustering, and classification. Decision trees, SVMs, and neural networks are examples.
Knowledge Representation Statistics
Linear regression, decision trees, and support vector machines help represent and analyse data. These models are useful for machine learning knowledge representation because they can capture complicated data linkages and patterns.
Probabilistic reasoning knowledge representation examples
Probabilistic medical diagnosis:
Imagine a patient with symptoms that may indicate Disease X and Disease Y. Multiple variables may affect symptom occurrence and intensity, although neither condition is exclusive. A doctor must diagnose accurately to decide therapy.
How to use probabilistic reasoning:
1. Symptom Observation: Doctors consider symptoms such as fever, tiredness, and joint discomfort.
The doctor use a probabilistic model or Bayesian network to depict the correlation between symptoms and illnesses. Given each disease's existence or absence, this model predicts certain symptoms.
3. Data Input: Doctors enter patient symptom data into the model. Disease X patients have a 70% risk of joint discomfort, whereas Disease Y patients have a 30% chance.
4. Probabilistic Inference: The model determines the conditional probability of Diseases X and Y based on observed symptoms. These estimates use previous probabilities (the population's illness base rate) and symptom likelihood.
5. Probabilistic diagnostic: The model delivers a diagnosis. It may suggest that the patient has 60% Disease X and 40% Disease Y.
6. Treatment Decision: Probabilistic diagnosis aids doctors in making educated treatment decisions. If Disease X is probable, they may pick a disease-specific therapy. They may also track the patient's development and reassess the diagnosis.
This example shows how probabilistic thinking helps doctors make more nuanced and data-driven judgements in difficult circumstances when symptoms may not indicate a diagnosis. It recognises medical diagnosis' ambiguity and helps patients obtain appropriate therapy when the evidence is inconclusive.

Here Is an Illustration of Knowledge Representation Through Statistical Reasoning.

Problem-solving examples employing statistical reasoning:

Example 1: Mean/SD

International Journal of Academic and Applied Research (IJAAR)
ISSN: 2643-9603
Vol. 8 Issue 2 February - 2024, Pages: 52-65

Calculate the average (mean) and measure of dispersion (standard deviation) for the dataset. {12, 15, 18, 22, 24, 27}.

Solution: 1. Calculate Mean (Average): Mean = (12+15+18+22+24+27) / 6 = 18.67 (rounded to two decimal places).

2. Calculate Standard Deviation: - Calculate squared differences from the mean for each data point:
(12 - 18.67)^2 = 44.89 (15 - 18.67)^2 = 13.56 (18 - 18.67)^2 = 0.45 (22 - 18.67)^2 = 11.09 (24 - 18.67)^2 = 30.69 (27 - 18.67)^2 = 69.16

- Calculate variance (average of squared differences): Variance = 24.63 (44.89 + 13.56 + 0.45 + 11.09 + 30.69 + 69.16) / 6.

Standard deviation (square root of variance): ≈ √(24.63) = 4.96 (rounded to two decimal places).

In this dataset, the calculated mean is 18.67, and the standard deviation is 4.96.

Probability Distribution Example 2:

Issue: Rolling a fair six-sided die to determine the probability of getting an even number.

When rolling a fair six-sided die, there are six equally probable outcomes: {1, 2, 3, 4, 5, 6}.

1. Identify positive results (even numbers): {2, 4, 6}.

2. Calculate the probability of rolling an even number:
   Probability = (Number of favorable outcomes) / (Total outcomes) = 3/6 = 1/2.

The probability of rolling an even number is 1/2.

Example 3: Hypothesis Testing

Problem: A manufacturer claims the typical light bulb lifetime is 1000 hours. A sample of 25 light bulbs has a mean of 980 hours and a standard deviation of 40 hours. Verify the manufacturer's claim at a 0.05 significance level.
Solution:
1. Create hypotheses:
Null Hypothesis (H0): Light bulbs have an average lifetime of 1000 hours ($\mu = 1000$).
- Alternative Hypothesis (H1): Light bulbs' average lifetime is not 1000 hours ($\mu \neq 1000$).

2. Set the significance threshold (α) to 0.05.

3. Formula for test statistic (t-score):
t = (Sample Mean - Population Mean) / (Sample Standard Deviation / √Sample Size).
t = (980 - 1000) / (40 / √25) = -2.5

4. Find the df for a t-distribution, which is 24 (sample size - 1).

5.

6. Compare the test statistic absolute value to the essential t-value:
$|t| = 2.5 > 2.064$

7. We reject the null hypothesis because the test statistic exceeds the threshold t-value.
8. Conclusion: Statistics show that light bulbs' average lifetime is likely not 1000 hours.

These examples show how statistical reasoning may compute means, probabilities, and test hypotheses to make data-driven choices.
Models Models of representation
Models simplify system or process details. Our emphasis is on knowledge representation models, but every model's

International Journal of Academic and Applied Research (IJAAR)
ISSN: 2643-9603
Vol. 8 Issue 2 February - 2024, Pages: 52-65

value relies on how effectively it reflects what it's meant to represent for its intended purpose. This depiction might be physical, metaphorical, or a mix. Form and abstraction levels vary among representations. A computer programmer's flowchart is a graphical model that illustrates time-based stages in space. A symbolic mathematical equation represents a system or problem's variable connections. Physical model is less abstract than schematic, which is less abstract than symbolic. More abstract models have less physical similarity to their subjects. Here are several models:

1. Physical Models: These fixed representations of objects, problems, or systems are easy to understand but lacking in value for decision analysis, except in limited applications. They fall into two categories: iconic and analogue models.
2. Symbolic Models: Utilising visual symbols, symbolic models represent changeable connections. icons symbolise them. They are abstract models. The symbolic model's main benefit is symbol manipulation. This makes them useful to decision-makers for analysis and prediction.
3. Mathematical models: Purpose and variables define mathematical models. Mathematical models may describe or prescribe. A descriptive model depicts reality, whereas a normative model suggests behaviour. Only systems or problems are described or solved by mathematical models. Descriptive models clarify a condition, indicate areas for change, and evaluate decision options. Descriptive models do not choose the optimal option. It gives a framework for analysis to help choose an option. A normative model determines the optimum option based on a decision criteria. The goal of optimising decision models is to find the best option under given circumstances.
4. Markov models: Andrey Markov's stochastic model, a Markov chain, predicts a succession of occurrences depending on the prior event's state. Finance and other sequential data businesses employ this simple approach. Even Google's page rank algorithm, which prioritises links, is a Markov chain. This model approximates future occurrences using our observations and mathematics.

The Markov process, often known as the Brand Switch model, determines the likelihood of switching states. One of Markov's main advantages is that a stochastic variable's future state depends solely on its current state. Stochastic variables are defined as variables whose values rely on random events.

The limited number of HMM states might be visible or invisible. Solid arrows link visible states, whereas dashed arrows connect unseen ones. As evidence is provided, the probability value of a state is shown at the top of the arrows. Figure 1 illustrates HMM.

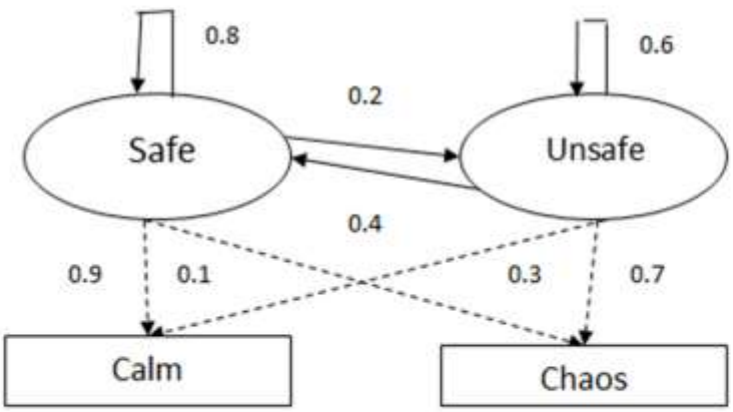


Figure 1 illustrates a Hidden Markov Model (HMM) describing the relationship between the states "Safe" and "Unsafe" with the observations (invisible states) "Calm" and "Chaos."

In the presented HMM:

a. Relating observable states to observable states:

- \(P(\text{Safe}|\text{Safe}) = 0.8 \)
- \(P(\text{Unsafe}|\text{Safe}) = 0.2 \)
- \(P(\text{Unsafe}|\text{Unsafe}) = 0.6 \)

International Journal of Academic and Applied Research (IJAAR)
ISSN: 2643-9603
Vol. 8 Issue 2 February - 2024, Pages: 52-65

- \(P(\text{Safe}|\text{Unsafe}) = 0.4 \)

b. Observable-nonobservable relationships:

- \(P(\text{Calm}|\text{Safe}) = 0.9 \)
- \(P(\text{Chaos}|\text{Safe}) = 0.1 \)
- \(P(\text{Calm}|\text{Unsafe}) = 0.3 \)
- \(P(\text{Chaos}|\text{Unsafe}) = 0.7 \)

Several applications of HMM are noted:

1. Bhusari & Pati (2011) utilized HMM for detecting credit card theft.
2. Gale & Young (2008) applied HMM in the field of voice recognition.
3. Hamdi & Frigui (2015) reported the use of HMM in landmine detection.
4. Hewahi (2010) suggested employing HMM for network management.
5. Naghizadeh et al. (2012) introduced a modified HMM for protein secondary structure prediction.

In 2013, Sherlock et al. developed linked HMM for illness interactions. ShivaPrasad & RaghuKisore (2015) used HMM to categorise metamorphic viruses. Tenyakov (2014) wrote a PhD thesis on HMM estimation and finance.
Some researchers improved HMM-representation systems. Particle swarm optimisation was used to construct new HMMs (Hewahi, 2015). (Hewahi, 2011a) advocated a neural evolution strategy to produce new HMMs, whereas (Hewahi, 2011b) proposed a genetic algorithm-only approach. New approaches for converting HMM to real-time filtered production rules are provided in (Hewahi, 2011c).

Hierarchical Hidden Markov Model

Fine, Singer, and Tishby (1998) presented a Hierarchical Hidden Markov Model (HHMM) to express connected states in several HMMs. Figure 2 shows HHMM.

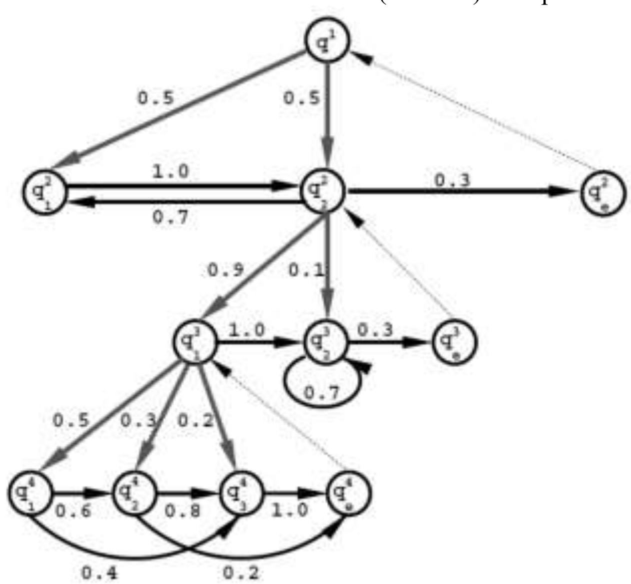


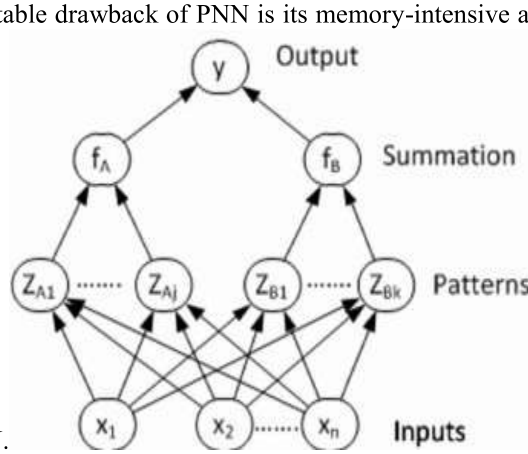
Figure 2. Four-level HHMM. Upper value is HMM level, lower value is state index, and q is state name.

International Journal of Academic and Applied Research (IJAAR)
ISSN: 2643-9603
Vol. 8 Issue 2 February - 2024, Pages: 52-65

4. Probability Neural Network

The Probability Neural Network (PNN) is a supervised learning feed-forward network characterized by four layers: input, pattern, summation, and output. In the first layer, nodes represent all input characteristics, while the second layer comprises nodes for each data set instance, with nodes having the same output being neighbors. The number of class label values determines the neurons in the third layer. The fourth layer consists of a single node producing the neural network output. This approach implements the statistical algorithm kernel discriminant analysis, as proposed by Specht (1992, 1998).

PNN provides several advantages, such as no need for training, absence of local minimum problems, no training required to add new instances to the dataset, and increased examples leading to more opportunities to find the best answer. However, a notable drawback of PNN is its memory-intensive and domain-specific nature. Figure 3 illustrates the structure of the PNN.



Application of Markov Model in a real-life situation

In Election

Consider sending popular politician AP to Congress. If AP has never been elected, their chances are 1/2.
• Losing candidates may run again the following year.
If currently in office, the AP has a 9/10 chance of being reelected. If they lose, they will quit politics.
To clarify this Markov Chain scenario, all statuses, transitions, and Markov property are fitted into 4 requirements:
Never elected, elected, continuing in office Retire
To create graphic 1, use shift curves to connect each status to its probability. All APs in this case quit without being elected.
This Markov Chain is also termed an Absorbing Markov Chain [11]. After entering absorptivity, one cannot depart. Each state of an absorbent Markov chain may absorb [12, 13]. An elected (E) AP has a 100% chance of being in office (C). In this case, the Markov property states that the probability distribution of any condition solely relies on its neighbours.
Except Different Markov Chain diagrams may be used to represent the simulation as shown in Figure 1. E and C may be merged as C1 since the probability of being elected (E) and presently in office (C) equals 1. Figure 2 shows the diagram. There are only three statuses in Figure 2.

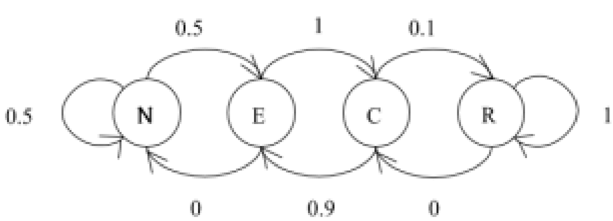


Figure 1. Markov Chain of voting Simulation within 4 status.

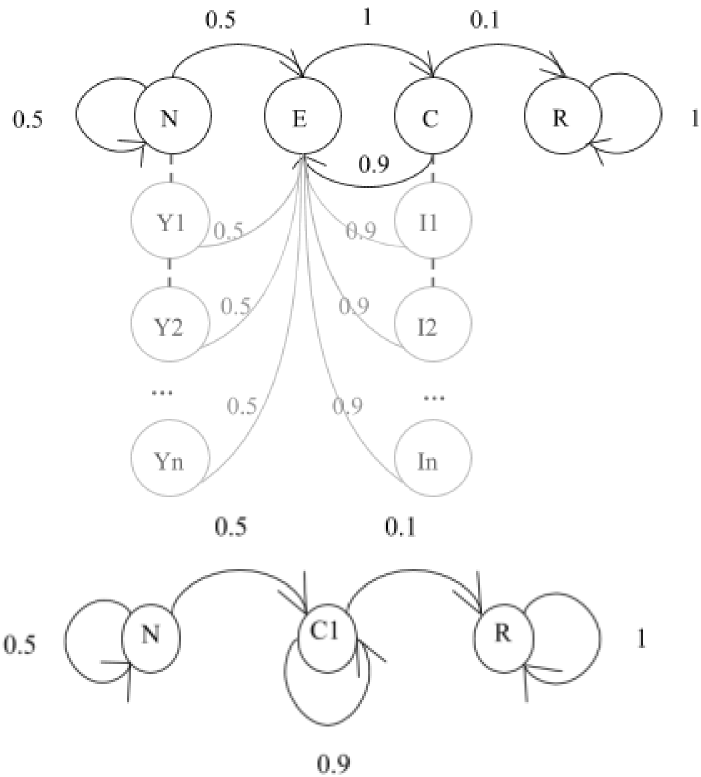


Figure 2. 3 state Markov chain voting simulation.

Two states seem to work, so consider splitting one into more. In fact, an AP may run for the first, second, or last time before retirement. Figure 3 illustrates that "Never Elected" (C) can be expanded to "First Year in Election" (Y1), "Second Year in Election" (Y2) and "New Year in Election" (Yn). "Currently in office" can also be expanded to "First year in office" (I1), "Second year in office" (I2) and "Nth year in office" (In) (Figure 3). The Markov property of the scenario does not rely on the frequency of the AP run, so the choice probability is the same every year (Yi & Ii). Because Figure 1 and Figure 3 calculate the same thing but have different numbers of states, this simulation has an unbounded upper bound on the states. **Figure 3.** Markov Chain of Voting Simulation with 4 status expansion.

Can the diagram be simplified now that there are several ways to graph this problem? Two-status graph. Never-elected and retirees may be out of office, therefore Figure 4 shows them as “Out of office” (O) and “In-office” (C1). It cannot show two probabilities in one curve. The AP may win if they have never been elected, but they will never return if they have retired. Only 3 statuses are needed for the voting simulation since two statuses with the same condition but different probability cannot be mixed.

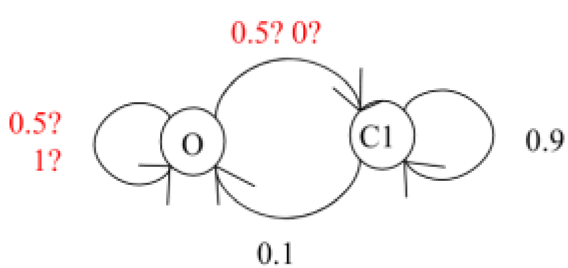


Figure 4. Two-Status Markov Chain Voting Simulation.

What is the lowest limit of statuses, and why may limitless statuses join one but not the other? Firstly, infinite statuses share probability. All probability distributions from expanded statuses to “election” (E) are the same, therefore they may be mixed. Now it's clear that the top and

International Journal of Academic and Applied Research (IJAAR)
ISSN: 2643-9603
Vol. 8 Issue 2 February - 2024, Pages: 52-65

bottom boundaries of the status number are more likely to predict based on probability from
other statuses than occurrences.

3. Application of Markov Chain in Weather Forecast

An average weather forecast predicts a 0.4 possibility of bright weather tomorrow, 0.3 probability of rainy or cloudy days, and a one-day duration for wet days and clouds.

• A wet day (R) followed by a cloudy day (C) is 0.4. A rainy day (R) followed by a sunny day (S) is 0.6. The likelihood of an overcast day becoming wet or sunny is 0.5. Since the weather prediction is a Markov Chain model, tomorrow's probability distribution relies only on today's weather. A tree diagram is in Figure 5.

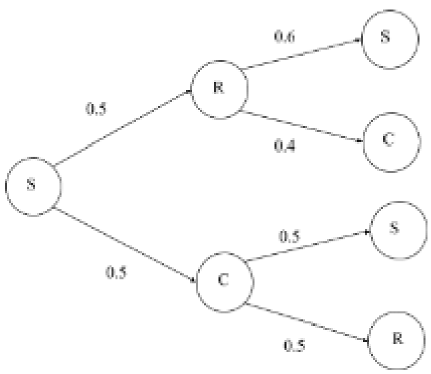


Figure 5. Weather forecast Markov Chain.

The three-event model seems difficult. According to the Markov property, Figure 6 may be condensed to Figure 7.

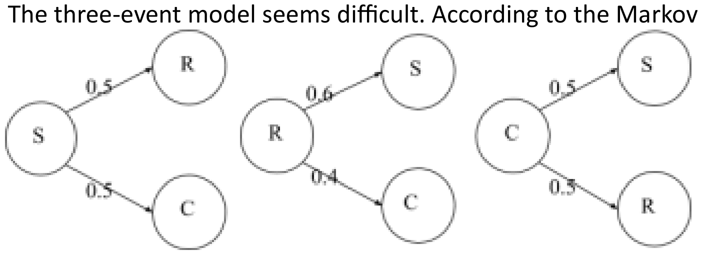
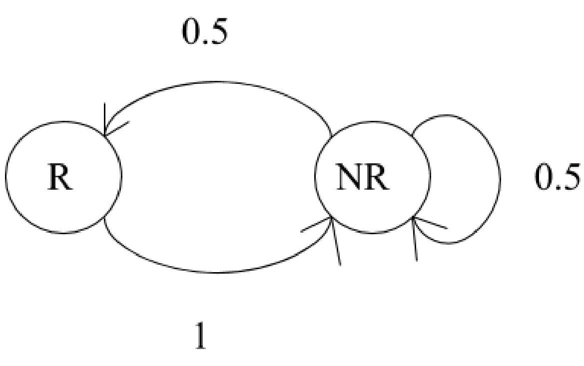


Figure 6. Weather Forecast Separation Markov Chain.

If To estimate the probable day of a rainy day in the week, the simulation may be represented by simply two statuses, "Rainy Day (R)" and "No Rain" (NR) (Figure 8), which incorporates bright and cloudy days. S and C followed by S or R have the same likelihood.

Sunshine and overcast days are not the same weather, yet in this situation, they may be combined since they have the same probability, proving the conclusion above.



next probability current	N	C1	R
N	0.5	0.5	0
C1	0	0.9	0.1
R	0	0	1

$$P = \begin{bmatrix} 0.5 & 0.5 & 0 \\ 0 & 0.9 & 0.1 \\ 0 & 0 & 1 \end{bmatrix}$$

International Journal of Academic and Applied Research (IJAAR)
ISSN: 2643-9603
Vol. 8 Issue 2 February - 2024, Pages: 52-65

1. Figure 7. Weather Markov ChainFigure 8: Weather Forecast Matrix…Simplified Forecast

In real life, if AP is running for the first time, how long should they retire? How likely are they to retire in 10 years? Matrix is a popular tool for solving Markov Chain issues, and its size depends on statuses. Multiply x with P till i = 46, then xi = [0.009 0.991].

99% of AP will retire in 47 years.

Within 10 years, the AP is 51.62% likely to resign (Figure 10).

Figure 9. Expect Matrix Calculation.Figure 10. Retirement Year Matrix Calculation Process.10-Year Retirement Probability

The most common kind of model is deterministic, having specific values for variables like pricing, expenses, and profits. We may need to predict future sales and presume a link. However, probabilistic modelling is a quantitative modelling method that projects a huge range of potential outcomes beyond recent history, takes into account new conditions and many uncertainties, and does not underestimate risks. Probabilistic index insurance models include variability and parameter uncertainty because they impact product performance forecasts. Probabilistic models need suitable probability distributions, precise input data, and proper accounting for variable linkages and relationships.

These probabilistic models are helpful in statistical analysis and have many good qualities. They make it easy to understand most data discrepancies. They may be developed hierarchically to generate complex models from simple pieces. Probabilistic modeling's innate protection against overfitting and coherent conclusions over various data types make it popular today.

Probabilistic Model Example

1. Linear Models

Generalised linear models use probabilistic modelling well. It greatly expands linear regression utilising exponential families. By combining observed values, simple linear regression predicts the anticipated return of an unknown element (the response variable, a random variable).

This indicates that each predictor change affects the response variable (linear response model). This is effective when the response variable fluctuates constantly in either direction or when any statistic, like human heights, changes by a small amount compared to the predictive parameters. Several response variables defy these assumptions.

Linear programming has solved various issues in many disciplines. However, commerce, industry, agriculture, and the military have used linear programming most. In many circumstances, management organisations must decide how to allocate limited resources. Money, labour, material, machine capacity, time, and technology are resources. Building a mathematical model is the next stage in addressing real-world issues. This involves three main steps:

1. Identifying solution variables (activity amount)
2. Create a linear objective function related to variables.
3. Identify system restrictions via linear connections of decision variables, reflecting limited resources in the issue.

Real-world linear programming examples

To maximise profit, a manufacturing business must identify the optimal number of each of three items given restricted resources. The labour and material needs and profit contribution for each product are as follows:

PRODUCT MIX PROBLEM				
RESOURCES	PRODUCT 1	PRODUCT 2	PRODUCT 3	AVAILABILITY
Labour(hr/unit)	5	2	4	240hr
Materials(lb/unit)	4	6	3	400lb
PROFIT (UNIT)	3	5	2	

Production has 240 hours of effort everyday. Supply materials are 400 pounds per say. The decision challenge is to manufacture how much of each product. For maximum profit. The issue fits all linear programming criteria.

Model the issue as linear programming.

Goal Variables

The problem's three choice variables are daily product 1, 2, 3 production. Represent these values symbolically as X1 = product quantity. 1. X2 = product amount 2. X3 equals product quantity 3.

Objective function

Product mix issue aims to maximise profit. As expected, overall profit is the sum of product profits. Product 1 profit is calculated by multiplying $3 by the number of units produced. Products 2 and 3 profit are calculated similarly. To get overall profit, use the formula: Maximise Z = 3x1 + 5x2 + 2x3, where 3x1 represents profit from product 1.

$2x3$ = profit from product 3 $5x2$ = profit from product 2

Limits on Systems

Production is restricted by labour and materials. Each of the three goods involves labour and resources. Product 1 takes 5 hours per unit to make. Thus, product 1 requires 5x1 hours of labour. Similar to product 2, product 3 requires 2x2 and 4x3 hours of labour. There are 240 production labour hours. Thus, labour constraint = $5x1 - 2x2 + 4x3 < 240$.

Material requirements are constrained similarly. Products 1 (x1), 3 (x2), and 3 (X3) require 6, 6, and 3 pounds, respectively. Given 400 pounds of raw materials, the constraint is $4X1 + 6X2 + 3X3 < 400$.

Since producing negative product quantities is illogical, we limit decision variable to a positive value. Mathematics explains nonnegativity constraints as $X1 > 0$, $X2 > 0$, $X3 > 0$. This inequality allows for flexibility in either case.

Complete linear programming can now be modelled mathematically:

Maximise $Z=3X1+5X2+2X3$.

Subject to $5X1 + 2X2 + 4X3 < 240$

$4X1 + 6X2 + 3X3 < 400$

$X1, X2, X3 > 0$

Challenges and Limitations

Statistical reasoning is the ability to use data, probability, and logic to conclude, make decisions and evaluate arguments. It is a valuable skill in many fields and contexts, but it also comes with some challenges and pitfalls that can affect its validity and usefulness.

1. Data quality and quantity: data quantity is concerned with the amount and diversity of the data. Poor data quality can lead to errors, biases, and misleading results; insufficient data quantity can limit the scope and generalizability of the analysis. Additionally, it is very important to check the validity, reliability, and credibility of your data sources and methods of collection.
2. Statistical literacy and communication: it is the ability to interpret, evaluate, and communicate statistical information, while statistical communication is the skill of presenting and explaining it to different audiences and purposes. Without proper, statistical literacy and communication, confusion, misunderstanding, or misuse of data analysis can occur. To improve your statistical literacy and communication, you should learn basic concepts and methods of statistics that relate to your field and context.
3. Cognitive biases and fallacies: this is to avoid or minimize the influence of cognitive biases and fallacies on your data and analysis. Cognitive biases are systematic errors in thinking that affect how we perceive, process, and remember information while fallacies are faulty or invalid arguments that violate the rules of logic or evidence. Both distort our statistical reasoning and lead to false or inaccurate conclusions, decisions, or arguments. To avoid such pitfalls, you should be aware of your own assumptions and preferences, and seek out different perspectives. Common cognitive biases and fallacies that can affect statistical reasoning include confirmation bias, selection bias, correlation-causation fallacy, survivorship bias, gambler’s fallacy, and appeal to authority fallacy.
5. Complexity: As problems and domains become more complex, the models used for statistical and probabilistic reasoning can become very intricate and computationally demanding. Maintaining and optimizing such models can be challenging.
6. Model Selection: Choosing the right statistical or probabilistic model for a specific problem is not always straightforward. Different models may be more or less appropriate for different scenarios, and selecting the wrong model can lead to suboptimal results.
7. Interpretability: Complex probabilistic models can be challenging to interpret. Understanding why a model makes a particular prediction or inference can be difficult, which can be a barrier to trust and acceptance, especially in critical applications like healthcare or autonomous systems.
8. Scalability: Scaling up probabilistic reasoning to handle large datasets or high-dimensional spaces can be computationally expensive. Efficient algorithms and distributed computing infrastructure are often required.
9.Assumptions and Independence: Many probabilistic models make assumptions about the independence of variables, which may not hold in real-world scenarios. Violation of these assumptions can lead to model inaccuracies.
10.Data Privacy and Security: Handling probabilistic data may raise privacy and security concerns, particularly in contexts where sensitive information is involved. Protecting data while still making meaningful inferences can be a delicate balance.
11.Training and Calibration: Training and calibrating probabilistic models can be challenging. Ensuring that the model's probabilities reflect the real-world likelihoods can require careful calibration and validation.
12.Overfitting and Underfitting: Like any machine learning approach, overfitting (fitting the model too closely to the training data) and underfitting (oversimplifying the model) can be issues in probabilistic modeling, requiring techniques like regularization to mitigate.
13.Domain Expertise: Effective use of statistical and probabilistic reasoning often requires domain expertise to design appropriate models, select relevant features, and interpret results. Integrating domain knowledge with statistical models can be a complex task.

Despite these challenges, statistical and probabilistic reasoning remains a critical approach in knowledge representation, particularly when dealing with uncertain, incomplete, or noisy data. Advances in machine learning and probabilistic modeling techniques continue to address many of these issues, making it possible to apply these methods to an expanding array of real-world problems and domains.

Conclusion knowledge representation using statistical and probabilistic reasoning has emerged as a powerful and versatile approach in the field of artificial intelligence and machine learning. This methodology provides a systematic way to encode and manipulate uncertain and complex information, enabling machines to make informed decisions, draw meaningful inferences, and adapt to various real-world scenarios.

Statistical and probabilistic reasoning allows us to model the inherent uncertainty and randomness in data and the real world. By leveraging techniques such as Bayesian networks, Markov models, and probabilistic graphical models, we can capture dependencies and relationships between variables, making it possible to handle real-world situations where traditional rule-based systems fall short. One of the significant advantages of this approach is its applicability in a wide range of domains, from natural language processing and computer vision to finance, healthcare, and autonomous systems. It enables machines to make rational decisions in uncertain environments and can even handle incomplete or noisy data gracefully.

However, it is important to acknowledge that knowledge representation using statistical and probabilistic reasoning is not without its challenges. It often requires substantial computational resources, especially for complex models and large datasets. Additionally, the quality of results heavily depends on the quality of data and the choice of appropriate models. Nonetheless, as technology advances and our understanding of statistical and probabilistic methods deepens, we can expect even more remarkable developments in this field. These techniques will continue to play a pivotal role in building intelligent systems that can reason, learn, and adapt in an ever-evolving world of uncertainty, making them invaluable tools for solving complex problems and enhancing our lives in numerous ways.

REFERENCES

Araghi, L., Khaloozade, H., & Arvan, M. (2009). *Ship Identification Using Probabilistic Neural Networks* (PNN), Proceedings of the International Multi Conference of Engineers and Computer Scientists 2009, Vol III, MECS 2009, Hong Kong, March 18 - 20.

Bhusari, V., & Pati, S. (2011). *Application of Hidden Markov Model in Credit Card Fraud Detection*, International Journal of Distributed and Parallel Systems, Vol.2, No.6, 203-2011.

El Emary, I., & Ramakrishnan, S. (2008). *On the Application of Various Probabilistic Neural Networks in Solving Different Pattern Classification Problems*, World Applied Sciences Journal, 4 (6), 772-780.

Fine, S., Singer, Y., & Tishby, N. (1998*). The Hierarchical Hidden Markov Model: Analysis and Applications,* in D. Haussler (Eds.), Machine Learning, 32, 41–62.

Gales, M., & Young, S. (2008). *The Application of Hidden Markov Models in Speech Recognition, Foundations and Trends in Signal Processing*, Vol. 1, No. 3,195-304.

Grim, J., Somol, P., & Pudil, P. (2005*). Probabilistic Neural Network Playing and Learning* Tic-*Tac-Toe, Pattern Recognition Letters-Special issue*: Artificial Neural Networks in Pattern Recognition, Vol. 26, issue 12,1866-1873.

Hamdi, A., & Frigui, H. (2015). *Ensemble Hidden Markov Models with Application to Landmine Detection,* EURASIP Journal on Advances in Signal Processing.

Hewahi, N. (2015). *Particle Swarm Optimization for Hidden Markov Model*, International Journal of Knowledge and Systems Science, Vol. 6, No. 2, 1-12.

Zhang, Q.. (2015*) Dynamic uncertain causality graph for knowledge representation and probabilistic reasoning:*. IEEE transactions on neural networks and learning systems, 26(7), 1503-1517.

Jain, D., Mosenlechner, L., & Beetz, M.. *2009 Equipping robot control programs with first-order probabilistic reasoning capabilities*. In *IEEE International Conference on Robotics and Automation* (pp. 3626-3631). IEEE.

Bellomarini, L., Benedetto, D., Laurenza, E., & Sallinger, E. (2022, August). *A framework for probabilistic reasoning on knowledge graphs*. In *International Conference on Soft Methods in Probability and Statistics* (pp. 48-56). Cham: Springer International Publishing.

Pearl, J. (1988) *Probabilistic reasoning in intelligent systems: networks of plausible inference.* Morgan kaufmann..

Benjamin, D. (2019). *probabilistic reasoning and judgment biases*. Handbook of Behavioral Economics: Applications and Foundations 1, 2, 69-186. .

Bacchus, F. (2016) *Representing and reasoning with probabilistic knowledge.*

Levesque, H. J. 2015 *Knowledge representation and reasoning*. Annual review of computer science, 1(1), 255-287.,