Building a Bayesian Network-Based Query System for Weather Prediction

by

Tian Zhiwen[1] (2230033036)

Yan Shan[1] (2230033048)

Li Hanfeng[2] (2230034025)

Hou Siqi[2] (2230034018)

Li Di[2] (2230034024)

Group 4 Project (Bayesian Networks AI3043)

Bachelor of Science (Honours)

in

[1]Applied Mathematics

[2]Artificial Intelligence

at

Beijing Normal University - Hong Kong Baptist University

May 2025

DECLARATION

We hereby declare that all the work done in this Project is of our independent effort. We also certify that we have never submitted the idea and product of this Project for academic or employment credits.

Contribution of Team Members

- **Tian Zhiwen**: Data processing, Bayesian network construction, coding and report writing for this part, overall typesetting of the report.
- **Yan Shan**: Data processing, Bayesian network construction, coding and report writing for this part.
- **Li Hanfeng**: Model Validation, coding and report writing for this part.
- **Hou Siqi**: Query System, Performance Evaluation, coding and report writing for this part.
- **Li Di**: Query System, Performance Evaluation, coding and report writing for this part.

All the members have contribution on topic determined, data collection and reference searching.

Tian Zhiwen (2230033036)	田质文
Yan Shan (2230033048)	[illegible]
Li Hanfeng (2230034025)	李涵锋
Hou Siqi (2230034018)	侯思齐
Li Di (2230034024)	李迪

Date: ____2025_05_03_____

Abstract

This study aims to predict the weather type based on a meteorological dataset. In this project, we build a weather predicting model based on Bayesian network. To achieve an accurate Bayesian model for classifying and predicting weather types, we first develop a framework to analyze and classify the original data. Then we apply `pgmpy` libraries to help us build the framework of the model, in which we use BIC to find the optimal structure, the `networkx` to visualize Bayesian network as DAG, MLE to determine conditional probability of each parameter, variable elimination and clique tree for calculation. After constructing the Bayesian network, we try to validate the network. For validation we use BIC and K2 and the results show that the values for BIC and K2 are extremely large in absolute value, which we believe it may due to the size of our dataset. For the end of our work, we evaluate the Bayesian network model through predictive consistency and inference efficiency. Using Variable Elimination and Clique Tree Belief Propagation, the model achieves identical classification accuracy '0.865' on a hold-out set. Efficiency experiments reveal two modes: Real-time and Batch mode with results of '0.73 ms/query' and '7.9 ms/query'. Additionally, an interactive query system powered by variable elimination and `Streamlit` enables users to compute posterior distribution. The system works with categorical and discretized numeric variables, and the results are shown as tables that can be easily visualized in bar charts.

Keyword: weather prediction, Bayesian network, MLE, VE, Clique Tree,BIC, K2, query system

Contents

1 Introduction

1.1 Background

Recent years, the combination of meteorological science with AI has opened new frontiers in weather forecasting research. Traditional methods like numerical weather prediction (NWP), which mainly rely on simulating atmospheric movements, have a limit on predicting localized extreme weather events. Unlike traditional models, Bayesian network integrates historical data with real-time observations through probabilistic modeling, enabling dynamic updates of predictions while quantifying uncertainties. For instance, Bayesian networks can reveal correlations among meteorological variables to develop a better prediction model which enhances the reliability of short-term weather forecasts.[1]

2 Our Dataset

2.1 Data Collection

The dataset used in this research is from `Kaggle` which contains eleven features. This dataset is generated by stimulating real weather data. The eleven features are as follows:

Table 1: Description of Variables in the Dataset

Original Variable	**Data Type**	**Description**
Visibility (km)	Numeric	Visibility in km, with extreme values.
Weather Type	Categorical	Target variable for weather classification.
Season	Categorical	Season of data recording.
Humidity ($\%$)	Numeric	Humidity percentage, may exceed 100%.
Atmospheric Pressure (hPa)	Numeric	Atmospheric pressure in hPa, wide - range.
Precipitation ($\%$)	Numeric	Precipitation percentage, with outliers.
UV Index	Numeric	Indicates ultraviolet radiation strength.
Wind Speed ($\mathrm{km\,h^{-1}}$)	Numeric	Wind speed in km/h, may have high outliers.
Location	Categorical	Type of data - recording location.
Temperature ($^\circ\mathrm{C}$)	Numeric	Celsius temperature, wide - range from cold to heat.
Cloud Cover	Categorical	Description of cloud cover.

2.2 Data Processing

Data preprocessing can ensure the reliability and consistency of the data. The way we predict the weather is to create several different predictions and each prediction shows the possible outcome of the weather.

First, we use 'data.isnull().sum()' to ensure that there is no missing value. Then, we present the distribution of features and their relationship with the target variable (weather type) through box plots and scatterplots. Finally, based on the applicability of Bayesian models to discrete data, the variables are classified into categorical variables and numerical variables. We use 'LabelEncoder' and 'KBinDiscretizer' seperately on two types of variables. Continuous data is divided into three intervals.

The encoding method adopts ordinal encoding, and the strategy is uniform division. The labels of each variable are mapped as follows:

Table 2: Variable Label Mappings

Variable	Value 1	Encoded as	Value 2	Encoded as	Value 3	Encoded as	Value 4	Encoded as
Season	Autumn	0	Spring	1	Summer	2	Winter	3
Location	coastal	0	inland	1	mountain	2		
Weather Type	Cloudy	0	Rainy	1	Snowy	2	Sunny	3
Cloud Cover	clear	0	cloudy	1	overcast	2	partly cloudy	3

3 Bayesian Network

3.1 Bayesian Framework Concept

The Bayesian method, grounded in Bayes' theorem $P(A|B) = \frac{P(B|A) \times P(A)}{P(B)}$, is applied in meteorological classification. Here, A denotes the weather type and B represents meteorological characteristics. The prior probability mirrors historical weather distribution, the likelihood function depicts the probability of observed data under specific weather, and the posterior probability, integrating historical and current data, offers the predicted probability of the weather type.

3.2 Bayesian Network

3.2.1 Notations

Table 3: Symbolic Notation of Variables

Original Variable	**Notation**	**Original Variable**	**Notation**
Visibility (km)	V	Precipitation (%)	P
Weather Type	WT	UV Index	UV
Season	S	Wind Speed ($\mathrm{km\,h^{-1}}$)	WS
Humidity (%)	H	Location	L
Atmospheric Pressure (hPa)	AP	Temperature (°C)	T
		Cloud Cover	CC

3.2.2 Structure Learning

The project employs a specific algorithm to learn the Bayesian network's structure, identifying potential connections among meteorological data features by continuously adjusting feature - connection relationships to find the optimal network.

We use the Bayesian Information Criterion (BIC), $BIC = -2 \times \ln(L) + k \times \ln(n)$ (explained in Part 4), as the evaluation metric. This criterion balances the model's fitting effect and complexity, effectively preventing overfitting.

To control the complexity of the model, the maximum in - degree of nodes is set to 4 and the maximum number of iterations is set to 100. The result edges in the **[Appendix A]**.

3.2.3 Construction

A discrete Bayesian network model is constructed after the above algorithm. This model takes the edges as the connection relationship to build a network that can reflect the probabilistic dependency relationship among meteorological features.

The `networkx` library is used to perform structural detection on the directed graph. We verify whether it is a directed acyclic graph through `nx.is_directed_acyclic_graph` and determine its weak connectivity by using `nx.is_connected`.

3.2.4 Visualization and Formula

Meteorological features are divided into five layers to construct a network structure, and the spacing of layers and nodes is set to determine the layout. The `networkx` and `matplotlib` libraries are used to generate the Bayesian network diagram for weather prediction.

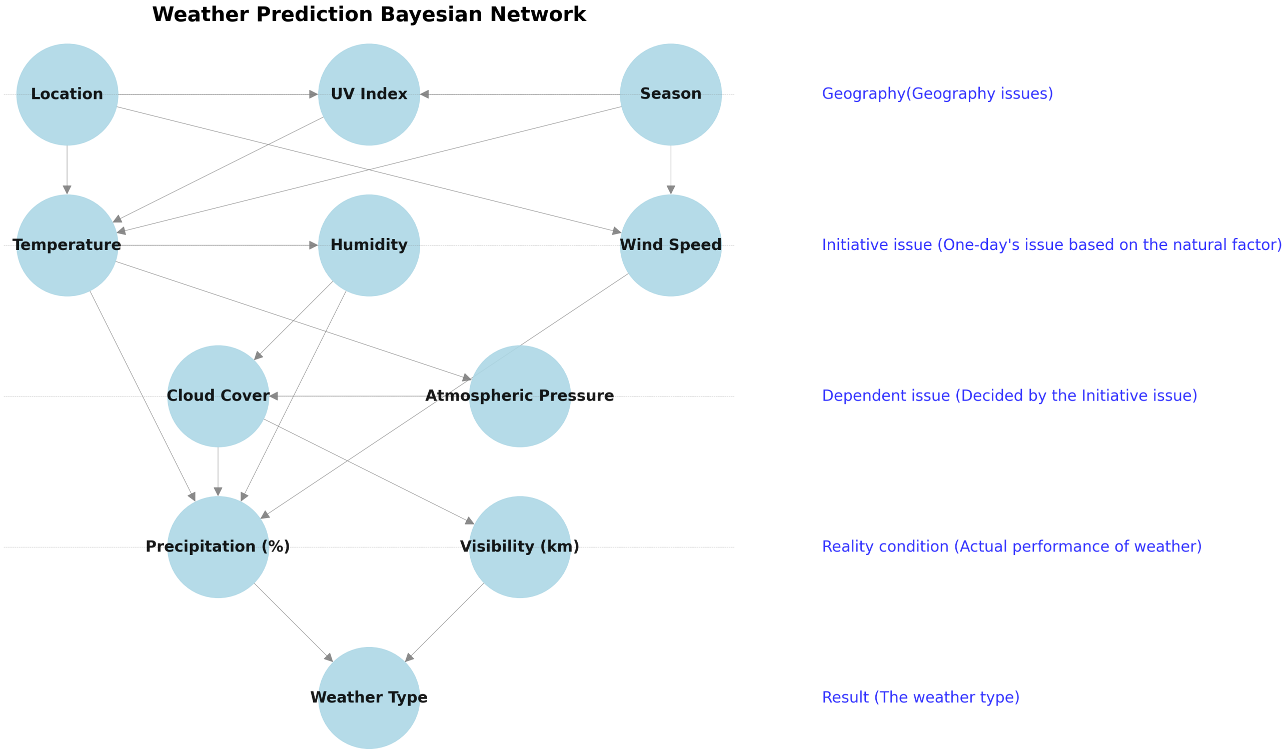


Figure 1: Weather Prediction Bayesian Network

By traversing the network edges, we construct the relationship between nodes and parent nodes. For each node, we output the conditional probability $P(\text{node}|\text{parents})$ or the prior probability $P(\text{node})$ respectively to clarify the probability dependency relationship of variables.

$$
\begin{aligned}
P(V, WT, S, H, AP, P, UV, WS, L, T, CC) = & \ P(V|CC)P(WT|V, P)P(S) \\
& \times P(H|T)P(AP|T)P(P|T, H, WS, CC) \\
& \times P(UV|L, S)P(WS|L, S)P(L) \\
& \times P(T|UV, S, L)P(CC|AP, H)
\end{aligned}
$$

3.3 Parameter Estimation and CPDs

By solving $\hat{\theta} = \arg\max_{\theta} \prod_{i=1}^{m} P(D_i|\theta)$ through maximum likelihood estimation (MLE), we determine the conditional probability distribution parameters of each node in the Bayesian network. We visually presents the probability dependencies among meteorological features and completes the meteorological data modeling and analysis. The CPDs in **[Appendix B]**.

3.4 Calculation of Meteorological Data Modeling

3.4.1 Variable Elimination

Variable elimination is used to get the probability for different conditions. We achieve meteorological prediction by eliminating the variables one by one in the joint probability distribution and simplifying with the help of the conditional probability formula and the probability chain rule. The probability chain rule is given by $P(X_1, X_2, \cdots, X_n) = \prod_{i=1}^{n} P(X_i|X_1, \cdots, X_{i-1})$.

3.4.2 Clique tree

To achieve high efficiency in meteorological data reasoning using Bayesian networks, the method of converting discrete Bayesian networks into a clique tree is adopted. Based on the principle of clique tree construction, the constructed discrete Bayesian network `Graph` is transformed into a clique tree `jt` through `graph.to_junction_tree()`. And the connections between these cliques should meet conditions such as probabilistic consistency.

After transformation and analysis, the clique tree contains the following cliques:

```
Clique 1: {P, CC, T, AP}
Clique 2: {P, CC, T, AP, L, S, WS}
Clique 3: {P, WT, V}
Clique 4: {P, CC, T, AP, L, S, V}
Clique 5: {UI, L, AP, T, S, V}
```

These cliques and their relationships form a clique tree structure, laying the foundation for the subsequent probabilistic reasoning of meteorological data.

4 Validation

4.1 Bayesian Network Information Criterion

Bayesian Network Information Criterion (BIC) is a metric evaluates the model's performance and complexity. Formula is shown below:

$$\mathrm{BIC} = -2\ln(L) + k\ln(n)$$

Where L means the Maximum Likelihood Expectation (MLE) in this model, symbolizing the fitting degree of it. k means the number of parameters, used to measure the complexity of model. n means number of samples that observed. With larger n, the penalty $k\ln(n)$ becomes larger, which helps to balance complexity and performance. We use BIC as our metric to build our Bayesian Network model. Lower BIC means better performance of our model.

4.2 K2

K2 is another important metric assuming a uniform prior (Dirichlet prior with parameters set to 1) over the parameters, evaluates the performance of Bayesian Network through the posterior probability. Formula is shown below:

$$\mathrm{Score}(B) = \sum_{i=1}^{m}\sum_{j=1}^{q_i}\left[\log\left(\frac{(r_i-1)!}{(N_{ij}+r_i-1)!}\prod_{k=1}^{r_i} N_{ijk}!\right)\right]$$

Where

- m is the number of nodes in Bayesian Network (e.g., $X_1, X_2, ...X_m$).
- q_i is the number of possible configurations of parent nodes of X_i, for example, if X_i has 2 parent nodes, one with 2 possible values and another with 3, then $q_i = 2 \times 3 = 6$.
- r_i is the number of possible values X_i can take, if X_i is binary, then $r_i = 2$.
- N_{ij} is the number of instances in dataset where parents nodes of X_i are in their j-th configuration.
- N_{ijk} is X_i in j-th configuration and X_i takes its k-th value. For example, if X_i is binary, $N_{ij1} + N_{ij2} = N_{ij}$.

K2 score decomposes the log-liklihood of network structure into contribution from each node X_i and its parents configuration.

For each node, it computes the term $\log\left(\frac{(r_i-1)!}{(N_{ij}+r_i-1)!}\prod_{k=1}^{r_i} N_{ijk}!\right)$ of observing data under a uniform Dirichlet prior. Dirichlet distribution is defined over the simplex of vectors that describe probabilities, represent prior beliefs about the parameters of a multinomial distribution. This term penalizes overfitting by favoring simpler structure when data is sparse.

Higher the K2 value, the better the performance and complexity balance of the model.

4.3 Validation Result

Due to the reason of size of our datasets, the BIC and K2 value can be extremely large in absolute value because they are accumulation of each sample.

5 Evaluation

This section summarises the key findings obtained from the latest round of experiments, focusing on two complementary aspects: *predictive consistency* and *inference efficiency*.

5.1 Predictive Consistency

Using two independent inference engines—Variable Elimination (VE) and Clique Tree Belief Propagation (BP)—the model achieves an identical classification accuracy of **0.865** on a randomly sampled hold-out set. The coincidence of accuracy values confirms that the learned conditional–probability tables yield stable posteriors regardless of the underlying inference algorithm. A visual inspection of the confusion matrix shows that all four weather classes are represented in a balanced manner, with no systematic bias towards or against any category. Furthermore, the per-class recall bar-plot reveals that each class attains a recall clustered around the same level, underscoring the model's ability to provide uniform coverage across the entire label space.

5.2 Inference Efficiency

Timing experiments highlight two complementary operating modes:

- **Real-time mode.** With VE the average latency per query is *sub-millisecond* (approximately $0.73\,\text{ms}$). This makes the network well suited for interactive or on-device deployments where response time is critical.
- **Batch mode.** Clique Tree inference requires an initial junction-tree construction, yet each subsequent query is resolved through local message passing. Although a single query takes about $7.9\,\text{ms}$, the cost amortises when a sequence of queries shares the same structure, yielding high throughput for large-scale or streaming scenarios.

A compact bar chart visually contrasts the two average timings, making the latency gap immediately apparent. The prediction result in the **[Appendix C]**.

6 Query System

This project runs an interactive Bayesian query for weather-forecast given an instance of evidence using the algorithm **Variable Elimination (VE)**. The user provides the instance of evidence (partial) through a **Streamlit** interface along with the variable of interest to compute the posterior distribution: $P(\text{Target} \mid \text{Evidence})$. The system works with categorical and discretized numeric variables, and the results are shown as tables that can be easily visualized in bar charts. Showing practical application of Bayesian Networks for Real-World Probabilistic Inference. **[Appendix D]**

References

[1] M. Waqas, U. Wannasingha Humphries, B. Chueasa, and A. Wangwongchai, "Artificial intelligence and numerical weather prediction models: A technical survey," *Natural Hazards Research*, vol. 10, no. 2, pp. 123–145, 2025. [Online]. Available: https://doi.org/10.1016/j.nhres.2024.11.004.

Appendices

Appendix A: Edges

```
('Temperature', 'Humidity'),
('Temperature', 'Precipitation (%)'),
('Temperature', 'Atmospheric Pressure'),
(Location', 'UV Index'),
('Season', 'UV Index'),
('UV Index', 'Temperature'),
('Location', 'Wind Speed'),
('Season', 'Wind Speed'),
('Humidity', 'Precipitation (%)'),
('Atmospheric Pressure', 'Cloud Cover'),
('Humidity', 'Cloud Cover'),
('Wind Speed', 'Precipitation (%)'),
('Cloud Cover', 'Visibility (km)'),
('Visibility (km)', 'Weather Type'),
('Precipitation (%)', 'Weather Type'),
('Season', 'Temperature'),
('Location', 'Temperature')
```

Appendix B: CPDs

Table 4: CPD for Location

$Location$	probability
$\{l^0\}$	0.270530
$\{l^1\}$	0.364848
$\{l^2\}$	0.364621

Table 5: CPD for Season

$Season$	probability
$\{s^0\}$	0.189394
$\{s^1\}$	0.196818
$\{s^2\}$	0.188788
$\{s^3\}$	0.425000

Table 6: CPD for UV Index

$Location, Season$	uv^0	uv^1	uv^2
$\{l^0, s^0\}$	0.595556	0.268889	0.135556
$\{l^0, s^1\}$	0.586712	0.261261	0.152027
$\{l^0, s^2\}$	0.568581	0.248535	0.182884
$\{l^0, s^3\}$	0.593548	0.247312	0.159140
$\{l^1, s^0\}$	0.580605	0.258186	0.161209
$\{l^1, s^1\}$	0.573923	0.273574	0.152503
$\{l^1, s^2\}$	0.579014	0.237674	0.183312
$\{l^1, s^3\}$	0.784570	0.125632	0.089798
$\{l^2, s^0\}$	0.557072	0.285360	0.157568
$\{l^2, s^1\}$	0.573443	0.252644	0.173913
$\{l^2, s^2\}$	0.582547	0.251179	0.166274
$\{l^2, s^3\}$	0.807626	0.113518	0.078856

Table 7: CPD for Wind Speed

$Location, Season$	ws^0	ws^1	ws^2
$\{l^0, s^0\}$	0.877778	0.112222	0.01
$\{l^0, s^1\}$	0.871622	0.120495	0.007883
$\{l^0, s^2\}$	0.874560	0.112544	0.012896
$\{l^0, s^3\}$	0.861290	0.129032	0.009677
$\{l^1, s^0\}$	0.870277	0.120907	0.008816
$\{l^1, s^1\}$	0.904540	0.076834	0.018626
$\{l^1, s^2\}$	0.867257	0.115044	0.017699
$\{l^1, s^3\}$	0.796374	0.183390	0.020236
$\{l^2, s^0\}$	0.864764	0.119107	0.016129
$\{l^2, s^1\}$	0.883666	0.108108	0.008226
$\{l^2, s^2\}$	0.884434	0.103774	0.011792
$\{l^2, s^3\}$	0.807192	0.172877	0.019931

Table 8: CPD for Temperature

$Location, Season, UV Index$	t^0	t^1	t^2
$\{l^0, s^0, uv^0\}$	0.378731	0.608209	0.013060
$\{l^0, s^0, uv^1\}$	0.103306	0.876033	0.020661
$\{l^0, s^0, uv^2\}$	0.229508	0.729508	0.040984
$\{l^0, s^1, uv^0\}$	0.364683	0.618042	0.017274
$\{l^0, s^1, uv^1\}$	0.133621	0.857759	0.008621
$\{l^0, s^1, uv^2\}$	0.244444	0.711111	0.044444
$\{l^0, s^2, uv^0\}$	0.379381	0.608247	0.012371
$\{l^0, s^2, uv^1\}$	0.127358	0.849057	0.023585
$\{l^0, s^2, uv^2\}$	0.134615	0.826923	0.038462
$\{l^0, s^3, uv^0\}$	0.420290	0.572464	0.007246
$\{l^0, s^3, uv^1\}$	0.121739	0.860870	0.017391
$\{l^0, s^3, uv^2\}$	0.222973	0.722973	0.054054
$\{l^1, s^0, uv^0\}$	0.386117	0.605206	0.008677
$\{l^1, s^0, uv^1\}$	0.136585	0.839024	0.024390
$\{l^1, s^0, uv^2\}$	0.164062	0.812500	0.023438
$\{l^1, s^1, uv^0\}$	0.395538	0.594320	0.010142
$\{l^1, s^1, uv^1\}$	0.114894	0.872340	0.012766
$\{l^1, s^1, uv^2\}$	0.206107	0.740458	0.053435
$\{l^1, s^2, uv^0\}$	0.401747	0.591703	0.006550
$\{l^1, s^2, uv^1\}$	0.122340	0.851064	0.026596
$\{l^1, s^2, uv^2\}$	0.206897	0.731034	0.062069
$\{l^1, s^3, uv^0\}$	0.834498	0.163890	0.001612
$\{l^1, s^3, uv^1\}$	0.345638	0.647651	0.006711
$\{l^1, s^3, uv^2\}$	0.511737	0.474178	0.014085
$\{l^2, s^0, uv^0\}$	0.405345	0.585746	0.008909
$\{l^2, s^0, uv^1\}$	0.108696	0.882609	0.008696
$\{l^2, s^0, uv^2\}$	0.251969	0.716535	0.031496
$\{l^2, s^1, uv^0\}$	0.399590	0.594262	0.006148
$\{l^2, s^1, uv^1\}$	0.111628	0.865116	0.023256
$\{l^2, s^1, uv^2\}$	0.222973	0.756757	0.020270
$\{l^2, s^2, uv^0\}$	0.408907	0.585020	0.006073
$\{l^2, s^2, uv^1\}$	0.103286	0.882629	0.014085
$\{l^2, s^2, uv^2\}$	0.163121	0.787234	0.049645
$\{l^2, s^3, uv^0\}$	0.845494	0.152361	0.002146
$\{l^2, s^3, uv^1\}$	0.335878	0.648855	0.015267
$\{l^2, s^3, uv^2\}$	0.472527	0.521978	0.005495

Table 9: CPD for Humidity

$Temperature$	h^0	h^1	h^2
$\{t^0\}$	0.05999	0.535244	0.404766
$\{t^1\}$	0.25508	0.538866	0.206054
$\{t^2\}$	0	0.271605	0.728395

Table 10: CPD for Atmospheric Pressure

$Temperature$	ap^0	ap^1	ap^2
$\{t^0\}$	0.039493	0.919347	0.041160
$\{t^1\}$	0.023447	0.952821	0.023732
$\{t^2\}$	0	1	0

Table 11: CPD for Cloud Cover

$AtmosphericPressure, Humidity$	cc^0	cc^1	cc^2	cc^3
$\{ap^0, h^0\}$	0	0.306931	0.386139	0.306931
$\{ap^0, h^1\}$	0	0.410000	0.295	0.295
$\{ap^0, h^2\}$	0.25	0.25	0.250	0.25
$\{ap^1, h^0\}$	0.615913	0.037779	0.037207	0.309101
$\{ap^1, h^1\}$	0.121536	0.008434	0.471235	0.398795
$\{ap^1, h^2\}$	0.064048	0.000000	0.656492	0.279460
$\{ap^2, h^0\}$	0	0.378641	0.296117	0.325243
$\{ap^2, h^1\}$	0	0.322115	0.355769	0.322115
$\{ap^2, h^2\}$	0.25	0.25	0.25	0.25

Table 12: CPD for Visibility (km)

$CloudCover$	v^0	v^1	v^2
$\{cc^0\}$	0.383824	0.612436	0.003740
$\{cc^1\}$	0.389294	0.257908	0.352798
$\{cc^2\}$	0.845977	0.131527	0.022496
$\{cc^3\}$	0.616447	0.351754	0.031798

Table 13: CPD for Precipitation (%)

$CloudCover, Humidity, Temperature, WindSpeed$	p^0	p^1	p^2
$\{cc^0, h^0, t^0, ws^0\}$	0.333333	0.333333	0.333333
$\{cc^0, h^0, t^0, ws^1\}$	0.333333	0.333333	0.333333
$\{cc^0, h^0, t^0, ws^2\}$	0.333333	0.333333	0.333333
$\{cc^0, h^0, t^1, ws^0\}$	1	0	0
$\{cc^0, h^0, t^1, ws^1\}$	0.333333	0.333333	0.333333
$\{cc^0, h^0, t^1, ws^2\}$	0.333333	0.333333	0.333333
$\{cc^0, h^0, t^2, ws^0\}$	0.333333	0.333333	0.333333
$\{cc^0, h^0, t^2, ws^1\}$	0.333333	0.333333	0.333333
$\{cc^0, h^0, t^2, ws^2\}$	0.333333	0.333333	0.333333
$\{cc^0, h^1, t^0, ws^0\}$	0	0.166667	0.833333
$\{cc^0, h^1, t^0, ws^1\}$	0.	0	1
$\{cc^0, h^1, t^0, ws^2\}$	0.333333	0.333333	0.333333
$\{cc^0, h^1, t^1, ws^0\}$	0.945623	0.003979	0.050398
$\{cc^0, h^1, t^1, ws^1\}$	0	0.058824	0.941176
$\{cc^0, h^1, t^1, ws^2\}$	0.333333	0.333333	0.333333
$\{cc^0, h^1, t^2, ws^0\}$	0	0.173913	0.826087
$\{cc^0, h^1, t^2, ws^1\}$	0	0	1
$\{cc^0, h^1, t^2, ws^2\}$	0.333333	0.333333	0.333333
$\{cc^0, h^2, t^0, ws^0\}$	0	0.041667	0.958333
$\{cc^0, h^2, t^0, ws^1\}$	0	0	1
$\{cc^0, h^2, t^0, ws^2\}$	0.333333	0.333333	0.333333
$\{cc^0, h^2, t^1, ws^0\}$	0	0.042254	0.957746
$\{cc^0, h^2, t^1, ws^1\}$	0	0.060606	0.939394
$\{cc^0, h^2, t^1, ws^2\}$	0.333333	0.333333	0.333333
$\{cc^0, h^2, t^2, ws^0\}$	0	0.133333	0.866667
$\{cc^0, h^2, t^2, ws^1\}$	0	0	1
$\{cc^0, h^2, t^2, ws^2\}$	0.333333	0.333333	0.333333
$\{cc^1, h^0, t^0, ws^0\}$	0.359375	0.257812	0.382812
$\{cc^1, h^0, t^0, ws^1\}$	0.333333	0.333333	0.333333
$\{cc^1, h^0, t^0, ws^2\}$	0.333333	0.333333	0.333333
$\{cc^1, h^0, t^1, ws^0\}$	0.371795	0.346154	0.282051
$\{cc^1, h^0, t^1, ws^1\}$	0.333333	0.333333	0.333333
$\{cc^1, h^0, t^1, ws^2\}$	0.333333	0.333333	0.333333
$\{cc^1, h^0, t^2, ws^0\}$	0.333333	0.333333	0.333333
$\{cc^1, h^0, t^2, ws^1\}$	0.333333	0.333333	0.333333
$\{cc^1, h^0, t^2, ws^2\}$	0.333333	0.333333	0.333333
$\{cc^1, h^1, t^0, ws^0\}$	0.346457	0.362205	0.291339
$\{cc^1, h^1, t^0, ws^1\}$	0.333333	0.333333	0.333333
$\{cc^1, h^1, t^0, ws^2\}$	0.333333	0.333333	0.333333
$\{cc^1, h^1, t^1, ws^0\}$	0.333333	0.410256	0.256410
$\{cc^1, h^1, t^1, ws^1\}$	0.333333	0.333333	0.333333
$\{cc^1, h^1, t^1, ws^2\}$	0.333333	0.333333	0.333333
$\{cc^1, h^1, t^2, ws^0\}$	0.333333	0.333333	0.333333
$\{cc^1, h^1, t^2, ws^1\}$	0.333333	0.333333	0.333333
$\{cc^1, h^1, t^2, ws^2\}$	0.333333	0.333333	0.333333

Table 14: CPD for Precipitation (%)(continue)

$CloudCover, Humidity, Temperature, WindSpeed$	p^0	p^1	p^2
$\{cc^1, h^2, t^0, ws^0\}$	0.333333	0.333333	0.333333
$\{cc^1, h^2, t^0, ws^1\}$	0.333333	0.333333	0.333333
$\{cc^1, h^2, t^0, ws^2\}$	0.333333	0.333333	0.333333
$\{cc^1, h^2, t^1, ws^0\}$	0.333333	0.333333	0.333333
$\{cc^1, h^2, t^1, ws^1\}$	0.333333	0.333333	0.333333
$\{cc^1, h^2, t^1, ws^2\}$	0.333333	0.333333	0.333333
$\{cc^1, h^2, t^2, ws^0\}$	0.333333	0.333333	0.333333
$\{cc^1, h^2, t^2, ws^1\}$	0.333333	0.333333	0.333333
$\{cc^1, h^2, t^2, ws^2\}$	0.333333	0.333333	0.333333
$\{cc^2, h^0, t^0, ws^0\}$	0.294118	0.361345	0.344538
$\{cc^2, h^0, t^0, ws^1\}$	0.333333	0.333333	0.333333
$\{cc^2, h^0, t^0, ws^2\}$	0.333333	0.333333	0.333333
$\{cc^2, h^0, t^1, ws^0\}$	0.200000	0.435294	0.364706
$\{cc^2, h^0, t^1, ws^1\}$	0.333333	0.333333	0.333333
$\{cc^2, h^0, t^1, ws^2\}$	0.333333	0.333333	0.333333
$\{cc^2, h^0, t^2, ws^0\}$	0.333333	0.333333	0.333333
$\{cc^2, h^0, t^2, ws^1\}$	0.333333	0.333333	0.333333
$\{cc^2, h^0, t^2, ws^2\}$	0.333333	0.333333	0.333333
$\{cc^2, h^1, t^0, ws^0\}$	0.187614	0.408015	0.404372
$\{cc^2, h^1, t^0, ws^1\}$	0	0.412186	0.587814
$\{cc^2, h^1, t^0, ws^2\}$	0	0.166667	0.833333
$\{cc^2, h^1, t^1, ws^0\}$	0.394943	0.386225	0.218832
$\{cc^2, h^1, t^1, ws^1\}$	0	0.341615	0.658385
$\{cc^2, h^1, t^1, ws^2\}$	0	0.111111	0.888889
$\{cc^2, h^1, t^2, ws^0\}$	0	0	1
$\{cc^2, h^1, t^2, ws^1\}$	0	0	1
$\{cc^2, h^1, t^2, ws^2\}$	0	0	1
$\{cc^2, h^2, t^0, ws^0\}$	0	0.411673	0.588327
$\{cc^2, h^2, t^0, ws^1\}$	0	0.330579	0.669421
$\{cc^2, h^2, t^0, ws^2\}$	0	0.087719	0.912281
$\{cc^2, h^2, t^1, ws^0\}$	0	0.394251	0.605749
$\{cc^2, h^2, t^1, ws^1\}$	0	0.272358	0.727642
$\{cc^2, h^2, t^1, ws^2\}$	0	0.131579	0.868421
$\{cc^2, h^2, t^2, ws^0\}$	0	0	1
$\{cc^2, h^2, t^2, ws^1\}$	0	0.222222	0.777778
$\{cc^2, h^2, t^2, ws^2\}$	0	0	1
$\{cc^3, h^0, t^0, ws^0\}$	0.274336	0.389381	0.336283
$\{cc^3, h^0, t^0, ws^1\}$	0.333333	0.333333	0.333333
$\{cc^3, h^0, t^0, ws^2\}$	0.333333	0.333333	0.333333
$\{cc^3, h^0, t^1, ws^0\}$	0.879496	0.059353	0.061151
$\{cc^3, h^0, t^1, ws^1\}$	0.333333	0.333333	0.333333
$\{cc^3, h^0, t^1, ws^2\}$	0.333333	0.333333	0.333333
$\{cc^3, h^0, t^2, ws^0\}$	0.333333	0.333333	0.333333
$\{cc^3, h^0, t^2, ws^1\}$	0.333333	0.333333	0.333333
$\{cc^3, h^0, t^2, ws^2\}$	0.333333	0.333333	0.333333

Table 15: CPD for Precipitation (%)(continue)

$CloudCover, Humidity, Temperature, WindSpeed$	p^0	p^1	p^2
$\{cc^3, h^1, t^0, ws^0\}$	0.422780	0.395753	0.181467
$\{cc^3, h^1, t^0, ws^1\}$	0	0.423913	0.576087
$\{cc^3, h^1, t^0, ws^2\}$	0	0.100000	0.900000
$\{cc^3, h^1, t^1, ws^0\}$	0.586343	0.284964	0.128693
$\{cc^3, h^1, t^1, ws^1\}$	0	0.320000	0.680000
$\{cc^3, h^1, t^1, ws^2\}$	0	0	1
$\{cc^3, h^1, t^2, ws^0\}$	0	0	1
$\{cc^3, h^1, t^2, ws^1\}$	0	0	1
$\{cc^3, h^1, t^2, ws^2\}$	0	0	1
$\{cc^3, h^2, t^0, ws^0\}$	0	0.373096	0.626904
$\{cc^3, h^2, t^0, ws^1\}$	0	0.269231	0.730769
$\{cc^3, h^2, t^0, ws^2\}$	0	0.080000	0.920000
$\{cc^3, h^2, t^1, ws^0\}$	0	0.272131	0.727869
$\{cc^3, h^2, t^1, ws^1\}$	0	0.189349	0.810651
$\{cc^3, h^2, t^1, ws^2\}$	0	0.133333	0.866667
$\{cc^3, h^2, t^2, ws^0\}$	0	0.095238	0.904762
$\{cc^3, h^2, t^2, ws^1\}$	0	0	1
$\{cc^3, h^2, t^2, ws^2\}$	0	0	1

Table 16: CPD for Weather Type

$Precipitation(\%), Visibility(km)$	wt^0	wt^1	wt^2	wt^3
$\{p^0, v^0\}$	0.439567	0.013113	0.022805	0.524515
$\{p^0, v^1\}$	0.364241	0.008496	0.006649	0.620613
$\{p^0, v^2\}$	0.195652	0.268116	0.333333	0.202899
$\{p^1, v^0\}$	0.140807	0.424091	0.415082	0.020020
$\{p^1, v^1\}$	0.823077	0.055385	0.050769	0.070769
$\{p^1, v^2\}$	0.263514	0.270270	0.216216	0.250000
$\{p^2, v^0\}$	0.071548	0.428810	0.433818	0.065824
$\{p^2, v^1\}$	0.403017	0.086207	0.060345	0.450431
$\{p^2, v^2\}$	0.221477	0.214765	0.268456	0.295302

Appendix C: Weather Prediction Result

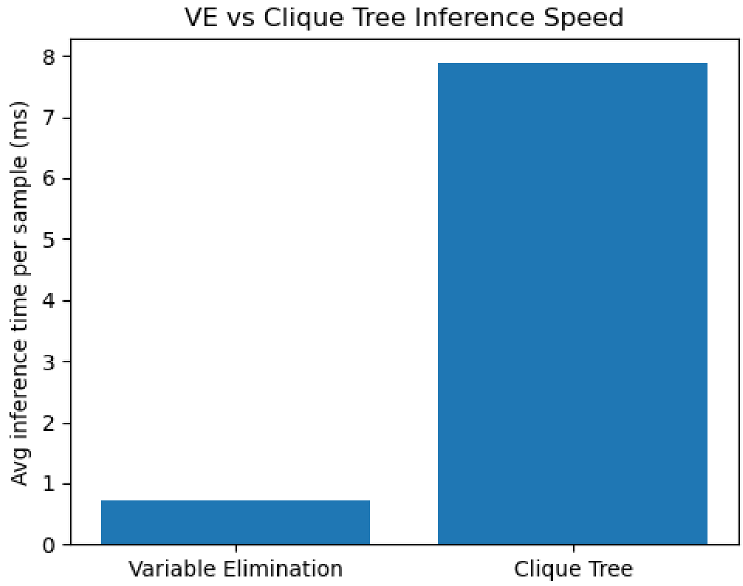


Figure 2: Inference Speed

	Method	Accuracy	Avg_time_ms
0	Variable Elimination	0.865	0.728232
1	Clique Tree	0.865	7.882838

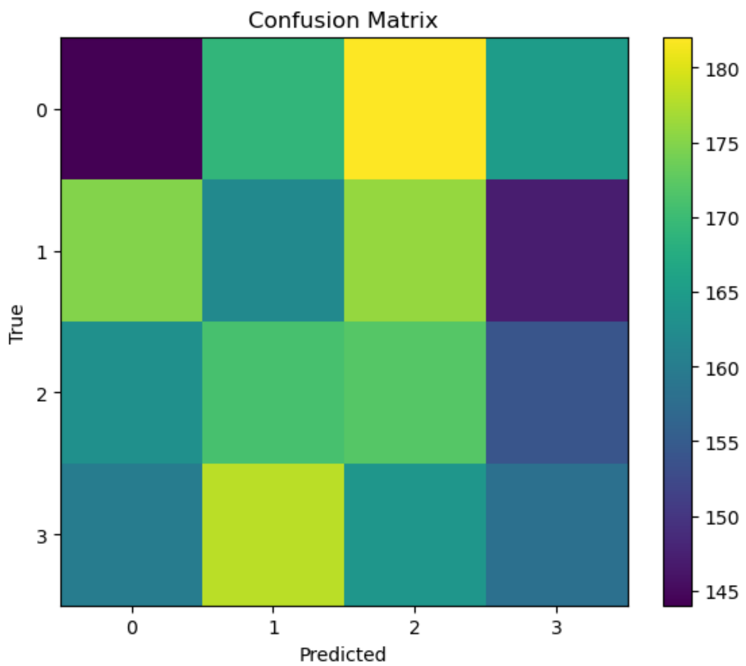


Figure 3: Confusion Matrix

Appendix D: Query System

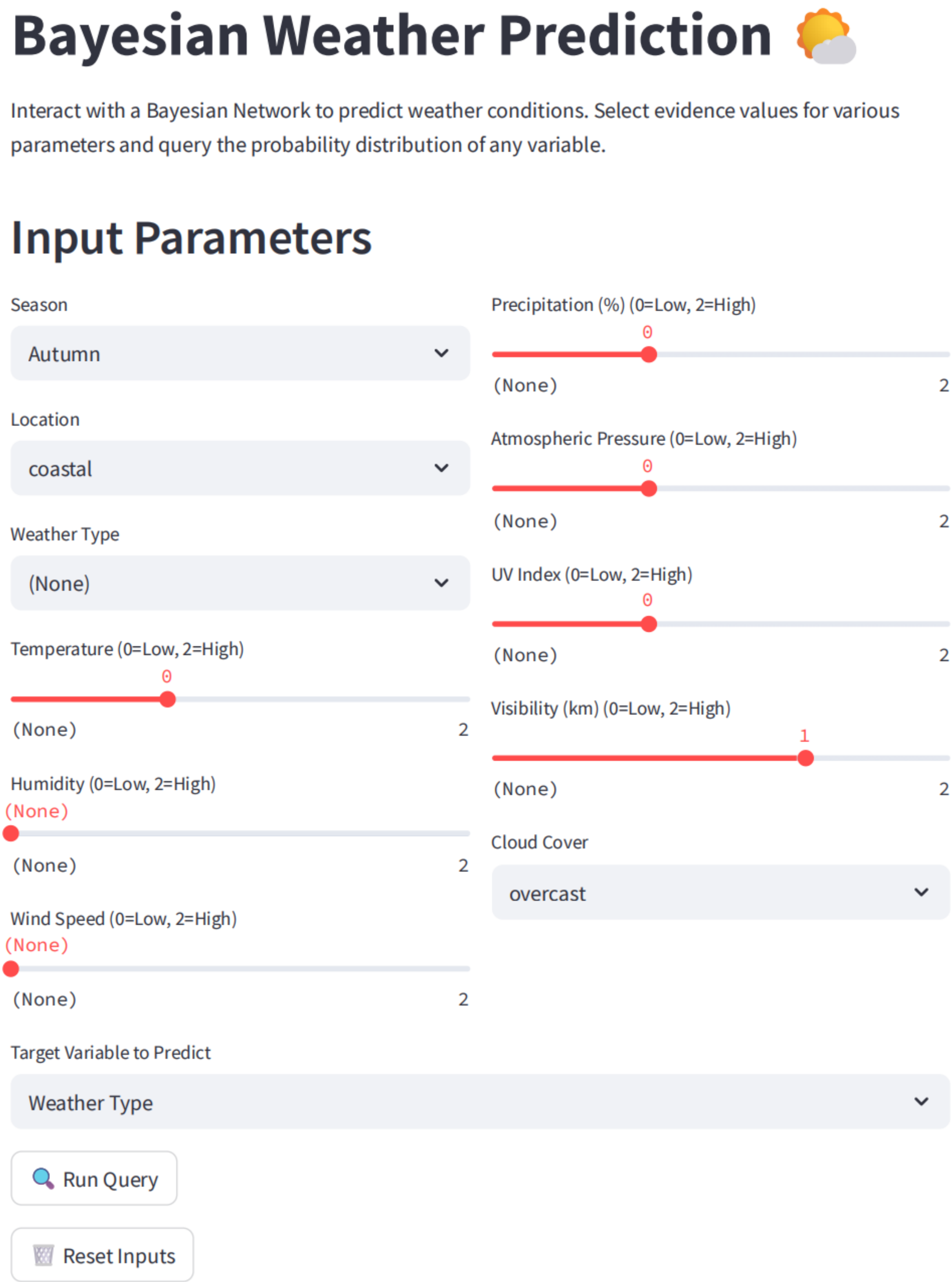
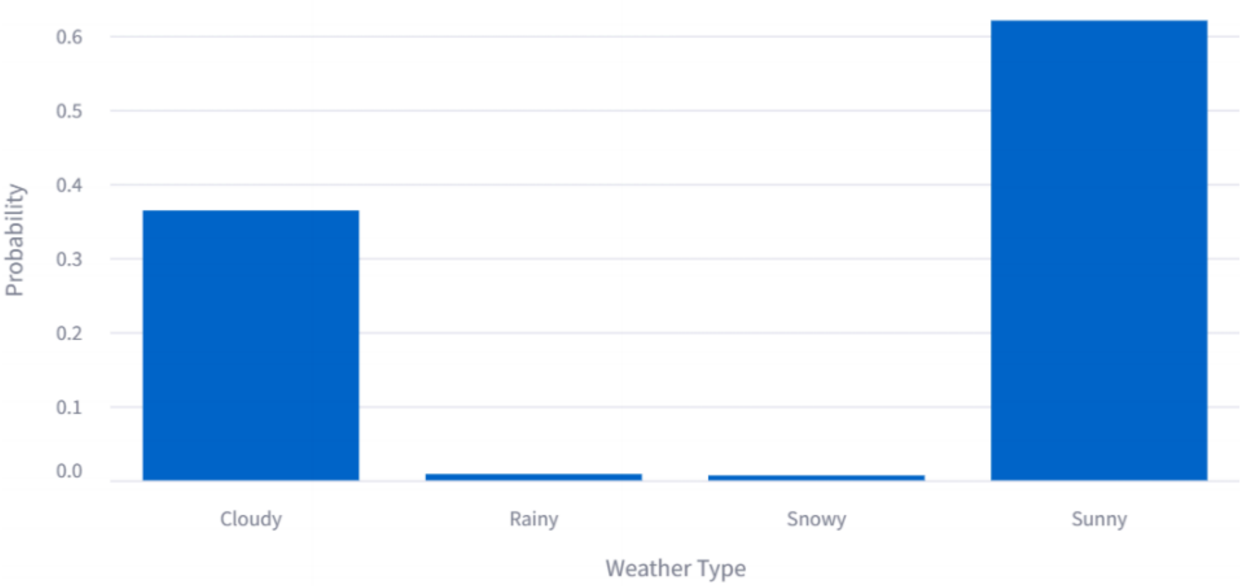


Figure 4

P(Weather Type | evidence)

	Weather Type	Probability
0	Cloudy	0.364
1	Rainy	0.008
2	Snowy	0.007
3	Sunny	0.621



Examples to Try:

- If Season = **Winter** and Humidity = **Low**, what's the Weather Type distribution?
- If Temperature = **High** and Cloud Cover = **Overcast**, is it likely Sunny?
- Combine Season = **Summer**, Weather Type = **Rainy**, and UV Index = **High** to see the result.

Figure 5