

How to Contextualize Empirical Data for Risk Analysis with LLMs: A Case Study of Power Outages

Haiyun Huang^{1*}, Yukun Li^{1*}, Marco A. Pretell¹,
Jacob Naroian¹, Ebadah Khan², Li-Ping Liu¹

¹ Tufts University ² Cornell University
{ haiyun.huang, yukun.li, liping.liu }@tufts.edu

Abstract

Large Language Models (LLMs) are increasingly being considered for high-stakes decision-making, yet their application in statistical risk analysis remains largely underexplored. A central challenge in this domain is enabling LLMs to effectively leverage historical data. To address this, we propose novel methods for extracting key information from raw data and translating it into structured contextual input within the LLM prompt. Applying our methods to a case study of power outage risk assessment, we demonstrate that this contextualization strategy significantly improves the LLM’s performance in risk assessment tasks. While the LLM’s prediction performance still does not match that of a standard machine learning model, the LLM-based approach offers distinct advantages in versatility and interpretability. These findings demonstrate a new paradigm for contextualizing data to support risk assessment.

1 Introduction

Large language models (LLMs) have been widely adopted across a broad range of domains due to their strong capabilities in understanding, generating, and reasoning over natural language (Naveed et al., 2025). However, LLMs face limitations in assessing the risks of adversarial events, as such assessments cannot rely solely on language-based reasoning and typically require grounding in empirical data (Haimès, 2011). Nevertheless, users often consult LLMs for preliminary risk assessments in order to obtain rapid and convenient guidance. This work studies methods for improving the risk assessment capabilities of LLMs, focusing on a case study of power outage risk assessment.

Rare but high-impact events require sustained attention to ensure preparedness. Accurate risk assessment for rare events depends on the availability of extensive historical data (Haimès, 2011). Power

outages provide a representative example: outages are typically sparse in both time and space, yet their consequences can be severe (Haimès, 2011). The causes of power outages are inherently complex, and precise prediction of their exact timing and location remains unfeasible. Nevertheless, robust statistical patterns can still be reliably identified within historical data.

Machine learning methods have been applied to predict power outages (Arora and Ceferino, 2023). By leveraging data from weather patterns, grid performance metrics, equipment health, and energy consumption trends, learning models are trained to identify patterns and predict outages before they occur, allowing for proactive maintenance and resource allocation. However, learning models are heavy tools that require expertise with depth and high-quality data. For example, a model often suffers from performance degradation if the input data is missing or erroneous in production environments (Emmanuel et al., 2021). A learning system, which requires a significant amount of effort to develop, could not provide quick answers to simple requests.

With the widespread adoption of LLMs, it is natural for users to turn to them for risk analysis (Cruz et al., 2024). Our preliminary studies indicate that LLMs can perform risk reasoning based on common knowledge. To be specific to a problem, risk analysis must be supported by historical data on the same problem.

A key research question is how to incorporate historical, region-specific data in a form that enables LLMs to conduct meaningful risk analysis. Naively including raw data in the prompt context is ineffective: the limited context window prevents inclusion of the full dataset, while a small number of examples fails to convey the necessary statistical structure for robust analysis. In this work, we propose three methods of contextualizing the data and describing the information with text inputs to the LLM. We test our methods in a case study of

*Joint first authors

power outage assessment and show that the proposed method can significantly improve the LLM’s risk assessment ability.

2 Related Work

Research on rare-event prediction emphasizes the need to extract statistical information from data to address the challenge of quantifying risks associated with rare adversarial events (Shyalika et al., 2024; Chen et al., 2024).

In the context of power outage prediction, machine learning has been widely applied to assess failure risks and restoration processes (Toubeau et al., 2022; Arora and Ceferino, 2023; Castillo, 2014). Prior studies combine meteorological, environmental, and infrastructure data to predict outages under extreme weather or equipment stress. Models range from logistic regression and ensemble learning (Eskandarpour and Khodaei, 2016; He and Cheng, 2021) to deep learning and probabilistic frameworks (Yang et al., 2020; Ghasemkhani et al., 2024; Gu et al., 2023; Zhai et al., 2021; Wang et al., 2024). Recent works have also used non-parametric and Bayesian methods for localized and non-weather-related outage prediction (Wedagedara et al., 2024; Owolabi and Sunter, 2022). The development of a learning system usually requires high-quality data and significant modeling effort.

Recently, LLMs have been applied to hazard assessment and product safety analysis (Collier et al., 2025; Kattamreddy and Chinnam, 2025; Diemert and Weber, 2023). In these applications, LLMs perform multi-step reasoning and enumerate plausible scenarios to identify risks arising from various conditions, such as design defects. LLMs are effective in such settings because the risks under consideration can largely be inferred through logical reasoning and domain knowledge. In contrast, the risk analysis studied here relies on statistical information derived from historical data rather than purely on reasoning. Consistent with this distinction, Cruz et al. (2024) examine risk assessment without access to data and show that LLM-generated risk scores lack accuracy.

3 Methods

3.1 Risk Measurement for Rare Events

We quantify the risk of power outages using an approach analogous to that employed in epidemiology. The *absolute risk* of rare events is typically

measured by their probability (Porta et al., 2014). Following standard terminology, we refer to rare events as *positive* cases and normal conditions as *negative* cases. Traditionally, the risk of positive events under a given condition c is defined as

$$R_c = \frac{N_c^{\text{pos}}}{N_c^{\text{pos}} + N_c^{\text{neg}}}.$$

Here N_c^{pos} and N_c^{neg} denote the number of positive and negative cases under c respectively. Unlike epidemiology, where the number of normal cases is typically easy to estimate, power outage datasets report only positive events. To address this limitation, we construct a finite negative sample set satisfying the following properties: (i) The number of negative events N_c^{neg} under each condition c remains proportional to the time–location extent over which the grid operates normally. (ii) For all c , the number of negative events in the sample is much greater than the positive events, $N_c^{\text{neg}} \gg N_c^{\text{pos}}$.

We use the *relative risk increase* to check the risk increase under various conditions. We approximate the conditional absolute risk by:

$$\tilde{R}_c = \frac{N_c^{\text{pos}}}{N_c^{\text{neg}}}$$

Similarly, the baseline absolute risk (covering all conditions) is approximated by:

$$\tilde{R}_{\text{base}} = \frac{N_{\text{base}}^{\text{pos}}}{N_{\text{base}}^{\text{neg}}}$$

Finally, the *relative risk increment* under the condition is defined as

$$\Delta R_c = \frac{\tilde{R}_c - \tilde{R}_{\text{base}}}{\tilde{R}_{\text{base}}} \quad (1)$$

Notably, the relative risk increment ΔR_c is robust to changes in the size of the negative sample, considering that the ratio of N_c^{neg} and $N_{\text{base}}^{\text{neg}}$ are preserved according to the condition.

Our aim is to use ΔR_c as a means of contextualizing the data and conveying risk information to the LLM in order to enhance its risk analysis capabilities.

3.2 Data Preparation and Analysis

In this study, we focus on quantifying the risk of power outages under specified conditions. To achieve robust performance, we integrate multiple data sources, including power outage reports, high-resolution weather forecasts, census data, and

Gust Interval (m/s)	Relative Risk Increase
0 – 9.5	≈ -0.5
9.5 – 11.3	-0.18
11.3 – 15.1	+0.16 – +1.01
15.1 – 18.9	+2.50 – +5.46
18.9 – 22.6	+12.6 – +25.1
22.6 – 28.3	+52 – +133
28.3+	+226 – ∞

Table 1: Relative risk increase across different gust speed intervals. Risk begins to rise noticeably above **11 m/s** and escalates rapidly beyond **18 m/s**.

spatial features, to model the probability of power outages across varying environmental and infrastructural contexts.

We obtain the power outage data from August 2018 through the end of 2023 from the Massachusetts state portal (mass.gov) (see Appendix A.0.1). Each power outage event includes the time and location of the event, as well as the duration of the outage. We drop the duration and treat all of them as positive cases.

Features/Conditions. For each event, we locate the corresponding weather condition from forecasts retrieved from the NOAA High-Resolution Rapid Refresh (HRRR) model. We extracted meteorological variables most relevant to outage occurrence, including wind, precipitation, temperature, etc. (see Appendix A.0.2). For each event, we also include census-based data such as population density and house density. We also include land cover characteristics such as cover type and tree canopy coverage (see Appendix A.0.3). Temporal factors reflect both meteorological and human activity patterns and play a significant role in power outages. To capture these effects, we also include month-of-year, day-of-lunar-cycle, day-of-week, and time-of-day into the features of each event.

Negative events. Negative events correspond to times and locations at which the power grid operates normally. Accordingly, negative events are assumed to be uniformly distributed over the same time range as positive events. Estimating the spatial coverage of the grid is challenging due to the lack of detailed power line data. As an approximation, we sample locations in proportion to housing density, which serves as a proxy for the spatial density of power infrastructure. To reduce sampling bias, locations are sampled independently for each negative event. We further exclude negative samples within 2 km of any recorded outage to avoid overlap with positive events. The total number of negative samples is set to ten times the number of

observed outage events.

Condition specification. We can specify a condition by imposing the range over feature values. Each condition slices out a subset of both positive and negative events, from which we can calculate the conditional risk \tilde{R}_c and then the relative risk increment according to Equation (1). For example, we can specify a range of the weather feature ‘‘Gust’’ and compute the risk because of gust winds. The risks at different gust speeds are shown in Table 1.

3.3 Risk analysis with LLMs

We use an LLM to perform the risk analysis based on the condition of an event. We convert the condition of each event (positive or negative) into a concise paragraph describing all relevant features, simulating a natural language report from a user about local conditions. Table 2 shows a few examples. Using a carefully designed prompt, the LLM is asked to assess the risk of a power outage for the described scenario and return its response in a structured JSON format with two fields: (i) a risk level (7 levels ranging from ‘‘very low’’ to ‘‘very high’’), and (ii) a brief reasoning statement explaining the factors contributing to the assessment.

For quantitative evaluation, each categorical risk level is mapped to a continuous scale (0–1) and compared with the ground-truth outage label, treating the task as a binary classification problem. We report the area under the ROC curve (AUC) and the F1 score for outage class as evaluation metrics.

As we mentioned above, the LLM can perform risk analysis with common knowledge. However, the common knowledge lack regional or quantitative specificity. The core problem is how to enable the LLM to use the information from the data to perform risk analysis.

Since it is neither feasible nor effective to input raw data directly to the LLM, we employ approaches to *contextualize* the data by extracting key information about risks and represent it in the text form. Specifically, we investigate the following three data contextualization methods.

- (i) **Risk Tabularization:** Tables showing each binned interval of a feature and the corresponding relative risk increase (Appx B.1);
- (ii) **Risk Summary:** A narrative summary of conditions that contributes significant relative risk increases. The summary is generated by GPT-4o based on tables from method (i). An example sentence is ‘‘Risk increases drastically

Feature Category	Example Features	Natural-Language Description Snippet
Temporal	Date = 2023-03-07, Day = Tuesday, Time = 14:28, Holiday = No, Lunar Cycle = 1.94 (Waxing Crescent)	“Today is March 7, Tuesday, and it is not a holiday. The time is 2 PM. The moon phase is waxing crescent.”
Meteorological	Temperature = 275.25 K, Dew Point = 269 K, Gust = 10.68 m/s, Wind = 6.90 m/s,	“It’s cold. The air is very dry. There are moderate wind gusts up to 23.9 mph.”
Spatial / Geographic	Town = Billerica, Latitude = 42.60, Longitude = -71.28, Utility = NationalGrid	“The location is in Billerica, Massachusetts. The utility group is NationalGrid.”
Socio-environmental	Population Density = 4.57 (\approx 457 people/km ²), Land Cover = Suburban mix, Tree Canopy = 48%	“The area has a low population density. The location is suburban, with houses, small shops, and moderate tree coverage.”

Table 2: A few examples showing the text description of event features.

with speed; GUST 17.5–20.3 \rightarrow risk \times 3.7” (Appx B.2);

- (iii) **Decision rules:** a subset of the most influential rules from a trained GBDT model. Each rule is converted into a sentence. An example rule is "IF Gust > 17.54m/s AND Gust <= 20.30m/s THEN Probability of power outage is 0.123" (Appx B.3).

These three methods condense a substantial amount of historical data into a highly compact textual form. The first two methods primarily capture information about single features, while the third is capable of representing low-order feature interactions.

4 Experiments

In this section, we present the results from our analysis with machine learning methods and LLMs. We evaluate three large language models, **Llama3-8b-Instruct** (Dubey et al., 2024), **Qwen-2.5-7B-Instruct** (Hui et al., 2024), and **GPT-4o** (Hurst et al., 2024), under four different contextual settings. We also tested using the GBDT, a traditional ML model, which serves as a non-LLM baseline. We hold out the data from 2023 as the test set, from which 450 events are randomly sampled to generate LLM queries. This setup enables direct comparison between the LLM-predicted risks and actual outage occurrences.

The results are shown in Table 3. We also provide one qualitative example in Appendix C. Across all three models, incorporating contextual information generally improves performance compared to the No-Context setting. GPT-4o demonstrates the most consistent gains, particularly under the Risk Summary Context, reflecting its superior ability to interpret semantically meaningful cues and integrate them into probabilistic reasoning.

Interestingly, Llama-3’s performance declines when given the full “Quantitative Risk” table but

Table 3: Comparison between LLMs under Different Context Settings

Context	Model	AUROC	F1
- No Context	Llama3-8b-Instruct	0.679	0.536
	Qwen-2.5-7B-Instruct	0.642	0.544
	GPT-4o	0.647	0.515
+ Quantitative Risk	Llama3-8b-Instruct	0.586	0.514
	Qwen-2.5-7B-Instruct	0.661	0.542
	GPT-4o	0.731	0.590
+ Risk Summary	Llama3-8b-Instruct	0.717	0.611
	Qwen-2.5-7B-Instruct	0.657	0.535
	GPT-4o	0.736	0.581
+ GBDT-derived Rule	Llama3-8b-Instruct	0.618	0.502
	Qwen-2.5-7B-Instruct	0.539	0.598
	GPT-4o	0.652	0.532
GBDT (Traditional ML Model)		0.907	0.728

improves when provided with the condensed “Risk Summary.” This suggests that Llama-3 struggles with dense numerical listings but benefits from simplified, linguistically structured representations such as “GUST > 28.3 \rightarrow risk \times 280.” In other words, LLMs appear more capable of leveraging compact, human-readable expressions of statistical patterns than exhaustive numeric data, indicating that interpretability and format simplicity can be as important as content richness when designing prompts for model reasoning.

Compared against the machine learning model, the LLM still has a lower accuracy even with the data information as the context. However, the responses of LLMs contains explanations of the assessment, making it easy to understand and also robust to missing and possibly erroneous features.

5 Conclusion and Future Work

This study shows that well-designed contexts incorporating statistical knowledge of the data can meaningfully improve the LLM’s performance in assessing the risk of power outages. This method can be easily extended into more applications.

Future work will focus on systematical explo-

ration of statistical and causal knowledge, as well as more adaptive prompt-engineering strategies tailored to model behavior. With deeper integration of data information and reasoning, LLMs hold promise as complementary tools for data-driven risk analysis in complex systems.

Limitations

The findings presented in this work are subject to the following limitations:

- (i) **Limited Model Generalizability:** Our evaluation focuses exclusively on a subset of available LLMs. Consequently, the observed performance and conclusions may not generalize to the broader landscape of large language models, particularly those not included in our assessment.
- (ii) **Scope of Risk Reasoning Evaluation:** We currently lack a dedicated dataset to definitively quantify the improvement in the LLM's risk reasoning ability resulting from the provided contextual information. While a risk assessment question-answering task could potentially demonstrate this improvement, resource constraints prevented the collection of such a specialized dataset.
- (iii) **Gap to Real-World Deployment:** While we successfully integrate contextual guidance, transitioning this work to a real-world operational setting would necessitate addressing critical practical considerations, including but not limited to safety validation, model interpretability, and robust human-in-the-loop procedures.

Acknowledgments

We thank the reviewers for their constructive feedback. Liu's work was supported by NSF 2239869. We utilized ChatGPT (OpenAI) for language refinement and programming assistance. All final content, results, and conclusions remain the sole contribution of the authors.

References

Prateek Arora and Luis Ceferino. 2023. Probabilistic and machine learning methods for uncertainty quantification in power outage prediction due to extreme events. *Natural Hazards and Earth System Sciences*, 23(5):1665–1683.

Anya Castillo. 2014. Risk analysis and management in power outage and restoration: A literature survey. *Electric Power Systems Research*, 107:9–15.

Wuxing Chen, Kaixiang Yang, Zhiwen Yu, Yifan Shi, and C. L. Philip Chen. 2024. A survey on imbalanced learning: latest research, applications and future directions. *Artif. Intell. Rev.*, 57:137.

Zachary A Collier, Richard J Gruss, and Alan S Abrahams. 2025. How good are large language models at product risk assessment? *Risk Analysis*, 45(4):766–789.

André F Cruz, Moritz Hardt, and Celestine Mendler-Dünner. 2024. Evaluating language models as risk scores. *Advances in Neural Information Processing Systems*, 37:97378–97407.

Simon Diemert and Jens H Weber. 2023. Can large language models assist in hazard analysis? In *International Conference on Computer Safety, Reliability, and Security*, pages 410–422. Springer.

Abhimanyu Dubey, Abhinav Jauhri, Abhinav Pandey, Abhishek Kadian, Ahmad Al-Dahle, Aiesha Letman, Akhil Mathur, Alan Schelten, Amy Yang, Angela Fan, and 1 others. 2024. The llama 3 herd of models. *arXiv e-prints*, pages arXiv-2407.

Tlameo Emmanuel, Thabiso Maupong, Dimane Mpoeleng, Thabo Semong, Banyatsang Mphago, and Oteng Tabona. 2021. A survey on missing data in machine learning. *Journal of Big data*, 8(1):140.

Rozhin Eskandarpour and Amin Khodaei. 2016. Machine learning based power grid outage prediction in response to extreme events. *IEEE Transactions on Power Systems*, 32(4):3315–3316.

Bitá Ghasemkhani, Recep Alp Kut, Reyat Yilmaz, Derya Birant, Yiğit Ahmet Arıkoç, Tugay Eren Güzeilyol, and Tuna Kut. 2024. Machine learning model development to predict power outage duration (pod): A case study for electric utilities. *Sensors (Basel, Switzerland)*, 24(13):4313.

Chan Gu, Chen Chen, and Wei Tang. 2023. Accurate and fast machine learning algorithm for systems outage prediction. *Solar Energy*, 251:286–294.

Yacov Y Haimes. 2011. *Risk modeling, assessment, and management*. John Wiley & Sons.

Jia He and Maggie X Cheng. 2021. Machine learning methods for power line outage identification. *The Electricity Journal*, 34(1):106885.

Binyuan Hui, Jian Yang, Zeyu Cui, Jiayi Yang, Dayiheng Liu, Lei Zhang, Tianyu Liu, Jiajun Zhang, Bowen Yu, Keming Lu, and 1 others. 2024. Qwen2. 5-coder technical report. *arXiv preprint arXiv:2409.12186*.

- Aaron Hurst, Adam Lerer, Adam P Goucher, Adam Perelman, Aditya Ramesh, Aidan Clark, AJ Ostrow, Akila Welihinda, Alan Hayes, Alec Radford, and 1 others. 2024. Gpt-4o system card. *arXiv preprint arXiv:2410.21276*.
- Ananth R Kattamreddy and Harisrujan Chinnam. 2025. The future of large language models in toxicological risk assessment: Opportunities and challenges. *Public Health Toxicology*, 5(1):1–3.
- Humza Naveed, Asad Ullah Khan, Shi Qiu, Muhammad Saqib, Saeed Anwar, Muhammad Usman, Naveed Akhtar, Nick Barnes, and Ajmal Mian. 2025. A comprehensive overview of large language models. *ACM Transactions on Intelligent Systems and Technology*, 16(5):1–72.
- Olukunle O Owolabi and Deborah A Sunter. 2022. Bayesian optimization and hierarchical forecasting of non-weather-related electric power outages. *Energies*, 15(6):1958.
- Miquel S Porta, Sander Greenland, Miguel Hernán, Isabel dos Santos Silva, and John M Last. 2014. *A dictionary of epidemiology*. Oxford university press.
- Chathurangi Shyalika, Ruwan Wickramarachchi, and Amit P. Sheth. 2024. A comprehensive survey on rare event prediction. *ACM Comput. Surv.*, 57(3).
- Jean-François Toubeau, Lorie Pardoën, Louis Hubert, Nicolas Marenne, Jonathan Sprooten, Zacharie De Grève, and François Vallée. 2022. Machine learning-assisted outage planning for maintenance activities in power systems with renewables. *Energy*, 238:121993.
- Xuesong Wang, Nina Fatehi, Caisheng Wang, and Masoud H Nazari. 2024. Deep learning-based weather-related power outage prediction with socio-economic and power infrastructure data. In *2024 IEEE Power & Energy Society General Meeting (PESGM)*, pages 1–5. IEEE.
- Harshana Wedagedara, Chandi Witharana, Robert Fahy, Diego Cerrai, Jason Parent, and Amal S Perera. 2024. Non-parametric machine learning modeling of tree-caused power outage risk to overhead distribution powerlines. *Applied Sciences*, 14(12):4991.
- Feifei Yang, David W Wanik, Diego Cerrai, Md Abul Ehsan Bhuiyan, and Emmanouil N Anagnostou. 2020. Quantifying uncertainty in machine learning-based power outage prediction model training: A tool for sustainable storm restoration. *Sustainability*, 12(4):1525.
- Chengwei Zhai, Thomas Ying-jeh Chen, Anna Grace White, and Seth David Guikema. 2021. Power outage prediction for natural hazards using synthetic power distribution systems. *Reliability Engineering & System Safety*, 208:107348.

A Data preparation

A.0.1 Power Outage Data: Positive Data

We collected outage records from the Massachusetts state portal (mass.gov), covering 2013–2023. Each record includes attributes such as outage duration, start and end times, and reported causes. In this study, we use only the event location (longitude, latitude) and the report time, which we use to define an outage (positive) event.

A.0.2 Weather Data

We obtained weather data from the High-Resolution Rapid Refresh (HRRR) model operated by the National Oceanic and Atmospheric Administration (NOAA). The HRRR model provides both real-time analysis and forecast data. However, we use only the forecast data in this work, as our goal is to predict potential outages before they occur—when only forecast information is available. The HRRR model issues forecasts every 6 hours, with predictions extending up to 48 hours ahead. For each specific outage event, we extract the forecast issued at 6:00 PM (UTC) on the previous day. We consider the following weather features, which are likely to be most relevant to power outages:

Variable Name	Level	Description
WEASD_1hr_acc_fcst	Surface	Water equivalent of accumulated snow depth (1-hr accumulated forecast)
WIND_1hr_max_fcst	10 m above ground	Maximum wind speed over the past hour
SHTFL	Surface	Sensible heat flux from the surface
PRATE	Surface	Precipitation rate (instantaneous)
DPT	2 m above ground	Dewpoint temperature
GUST	Surface	Wind gust at the surface
TMP	2 m above ground	Air temperature

Table 4: Weather features extracted from HRRR forecasts

A.0.3 Spatial Features: Geospatial and Census Data

Three datasets from the Massachusetts government were used for geospatial analysis.

Population density. Census information was obtained from MassGIS 2020 U.S. Census data. Although the dataset does not directly provide population density, it includes population counts for each census block (a tract of land) together with the block’s square footage. From these values, we computed population density at the block level, providing an estimate of population density across Massachusetts (pop20_density).

Utility service areas. Utility company service territories were obtained from MassGIS Public Utility Service Providers data. This dataset was used to assign the serving utility for each municipality (utility_group).

Land cover. Land cover data was retrieved from MassGIS 2016 Land Cover/Land Use data. This dataset contains 19 land cover categories, from which we derived two types of features: (i) level of land development, and (ii) type of land cover. The derived variables include: lc_dev_0, lc_dev_1, lc_dev_2, lc_dev_3, lc_is_water/wetland, lc_is_forest/shrub, lc_is_grass/crop, and tcc.

Data sources:

- [MassGIS Data: 2020 U.S. Census](#)
- [MassGIS Data: Public Utility Service Providers](#)
- [MassGIS Data: 2016 Land Cover/Land Use](#)

B Contexts

B.1 Quantitative Risk Context

The table below shows the relative risk (compared to the average) across different intervals for each analyzed feature.:

--- Relative risk and risk increase for Gust ---

	relative_risk	relative_risk_increase
feature_bin		
(-0.00509, 1.92]m/s	0.405	-0.595
(1.92, 3.8]m/s	0.425	-0.575
(3.8, 5.69]m/s	0.453	-0.547
(5.69, 7.57]m/s	0.523	-0.477
(7.57, 9.46]m/s	0.623	-0.377
(9.46, 11.3]m/s	0.819	-0.181
(11.3, 13.2]m/s	1.156	0.156
(13.2, 15.1]m/s	2.013	1.013
(15.1, 17]m/s	3.497	2.497
(17, 18.9]m/s	6.460	5.460
(18.9, 20.8]m/s	13.569	12.569
(20.8, 22.6]m/s	26.134	25.134
(22.6, 24.5]m/s	53.337	52.337
(24.5, 26.4]m/s	81.513	80.513
(26.4, 28.3]m/s	133.875	132.875
(28.3, 30.2]m/s	280.400	279.400
(30.2, 32.1]m/s	227.667	226.667
(32.1, 34]m/s	248.000	247.000
(34, 35.8]m/s	73.000	72.000
(35.8, 37.7]m/s	inf	inf

--- Relative risk and risk increase for Dew point Temperature ---

	relative_risk	relative_risk_increase
feature_bin		
(241, 244]K	0.091	-0.909
(244, 247]K	0.154	-0.846
(247, 250]K	2.655	1.655
(250, 253]K	1.042	0.042
(253, 255]K	0.624	-0.376
(255, 258]K	0.514	-0.486
(258, 261]K	0.613	-0.387
(261, 264]K	0.772	-0.228
(264, 267]K	0.717	-0.283
(267, 270]K	0.540	-0.460
(270, 273]K	0.642	-0.358
(273, 276]K	0.873	-0.127
(276, 279]K	0.772	-0.228
(279, 281]K	1.227	0.227
(281, 284]K	1.283	0.283
(284, 287]K	1.437	0.437
(287, 290]K	0.946	-0.054
(290, 293]K	1.171	0.171
(293, 296]K	1.657	0.657
(296, 299]K	1.409	0.409

--- Relative risk and risk increase for Percipitation rate ---

	relative_risk	relative_risk_increase
feature_bin		
(-0.001, 0.00049]kg/m^2/s	0.875	-0.125
(0.00049, 0.00098]kg/m^2/s	5.307	4.307
(0.00098, 0.00147]kg/m^2/s	8.675	7.675
(0.00147, 0.00196]kg/m^2/s	8.669	7.669
(0.00196, 0.00245]kg/m^2/s	9.022	8.022
(0.00245, 0.00294]kg/m^2/s	10.620	9.620
(0.00294, 0.00343]kg/m^2/s	15.156	14.156
(0.00343, 0.00392]kg/m^2/s	12.043	11.043
(0.00392, 0.00441]kg/m^2/s	9.714	8.714
(0.00441, 0.0049]kg/m^2/s	20.308	19.308

--- Relative risk and risk increase for Temperature ---

	relative_risk	relative_risk_increase
feature_bin		
(246, 249]K	0.000	-1.000
(249, 252]K	0.273	-0.727
(252, 256]K	0.936	-0.064
(256, 259]K	0.877	-0.123
(259, 262]K	0.831	-0.169
(262, 266]K	0.674	-0.326
(266, 269]K	0.670	-0.330
(269, 272]K	0.809	-0.191
(272, 276]K	0.691	-0.309
(276, 279]K	0.659	-0.341
(279, 282]K	0.987	-0.013
(282, 286]K	1.540	0.540
(286, 289]K	1.545	0.545
(289, 292]K	0.871	-0.129
(292, 296]K	0.992	-0.008
(296, 299]K	1.066	0.066
(299, 302]K	1.134	0.134
(302, 306]K	1.413	0.413
(306, 309]K	1.659	0.659
(309, 312]K	1.036	0.036

--- Relative risk and risk increase for popluation density ---

	relative_risk	relative_risk_increase
feature_bin		
(-12.4, 619]persons/km^2	1.934	0.934
(619, 1.24e+03]persons/km^2	0.706	-0.294
(1.24e+03, 1.86e+03]persons/km^2	0.510	-0.490
(1.86e+03, 2.47e+03]persons/km^2	0.428	-0.572
(2.47e+03, 3.09e+03]persons/km^2	0.327	-0.673
(3.09e+03, 3.71e+03]persons/km^2	0.276	-0.724
(3.71e+03, 4.33e+03]persons/km^2	0.308	-0.692
(4.33e+03, 4.95e+03]persons/km^2	0.321	-0.679
(4.95e+03, 5.57e+03]persons/km^2	0.261	-0.739
(5.57e+03, 6.19e+03]persons/km^2	0.388	-0.612
(6.19e+03, 6.8e+03]persons/km^2	0.257	-0.743
(6.8e+03, 7.42e+03]persons/km^2	0.587	-0.413
(7.42e+03, 8.04e+03]persons/km^2	0.389	-0.611
(8.04e+03, 8.66e+03]persons/km^2	0.346	-0.654
(8.66e+03, 9.28e+03]persons/km^2	0.275	-0.725
(9.28e+03, 9.9e+03]persons/km^2	0.695	-0.305
(9.9e+03, 1.05e+04]persons/km^2	0.847	-0.153
(1.05e+04, 1.11e+04]persons/km^2	0.268	-0.732
(1.11e+04, 1.18e+04]persons/km^2	0.127	-0.873
(1.18e+04, 1.24e+04]persons/km^2	0.300	-0.700

--- Relative risk and risk increase for day of month ---

	relative_risk	relative_risk_increase
Jan	0.532	-0.468
Feb	1.130	0.130
Mar	0.535	-0.465
Apr	0.932	-0.068
May	0.469	-0.531
Jun	0.672	-0.328
Jul	1.072	0.072
Aug	1.164	0.164
Sep	0.607	-0.393
Oct	2.914	1.914
Nov	0.785	-0.215
Dec	0.896	-0.104

--- Relative risk and risk increase for hour of day ---

	relative_risk	relative_risk_increase
Hour 0	0.625	-0.375
Hour 1	0.620	-0.380
Hour 2	0.665	-0.335
Hour 3	0.577	-0.423
Hour 4	0.649	-0.351
Hour 5	0.762	-0.238

Hour 6	0.926	-0.074
Hour 7	1.030	0.030
Hour 8	1.064	0.064
Hour 9	1.084	0.084
Hour 10	1.055	0.055
Hour 11	1.120	0.120
Hour 12	1.212	0.212
Hour 13	1.204	0.204
Hour 14	1.261	0.261
Hour 15	1.442	0.442
Hour 16	1.582	0.582
Hour 17	1.671	0.671
Hour 18	1.275	0.275
Hour 19	0.959	-0.041
Hour 20	0.829	-0.171
Hour 21	0.826	-0.174
Hour 22	0.771	-0.229
Hour 23	0.801	-0.199

--- Relative risk and risk increase for day of week ---

	relative_risk	relative_risk_increase
Sunday	0.568	-0.432
Monday	1.036	0.036
Tuesday	1.036	0.036
Wednesday	1.272	0.272
Thursday	1.087	0.087
Friday	1.158	0.158
Saturday	0.848	-0.152

--- Relative risk and risk increase for day of lunar cycle ---

	relative_risk	relative_risk_increase
Day 1	1.118	0.118
Day 2	1.737	0.737
Day 3	0.936	-0.064
Day 4	0.830	-0.170
Day 5	1.309	0.309
Day 6	2.196	1.196
Day 7	2.981	1.981
Day 8	3.381	2.381
Day 9	1.323	0.323
Day 10	1.068	0.068
Day 11	0.579	-0.421
Day 12	0.784	-0.216
Day 13	0.498	-0.502
Day 14	0.625	-0.375
Day 15	0.590	-0.410
Day 16	0.648	-0.352
Day 17	0.858	-0.142
Day 18	0.459	-0.541
Day 19	0.622	-0.378
Day 20	1.082	0.082
Day 21	0.802	-0.198
Day 22	0.481	-0.519
Day 23	0.670	-0.330
Day 24	0.693	-0.307
Day 25	0.578	-0.422
Day 26	0.511	-0.489
Day 27	0.650	-0.350
Day 28	0.444	-0.556
Day 29	0.892	-0.108

--- Relative risk and risk increase for tree canopy coverage ---

feature_bin	relative_risk	relative_risk_increase
(-0.1, 4.95]%	0.742	-0.258
(4.95, 9.9]%	0.608	-0.392
(9.9, 14.8]%	0.620	-0.380
(14.8, 19.8]%	0.649	-0.351
(19.8, 24.8]%	0.926	-0.074
(24.8, 29.7]%	1.196	0.196

(29.7, 34.6]%	1.134	0.134
(34.6, 39.6]%	1.267	0.267
(39.6, 44.5]%	1.255	0.255
(44.5, 49.5]%	1.276	0.276
(49.5, 54.5]%	1.538	0.538
(54.5, 59.4]%	0.917	-0.083
(59.4, 64.3]%	5.794	4.794
(64.3, 69.3]%	1.248	0.248
(69.3, 74.2]%	1.300	0.300
(74.2, 79.2]%	0.908	-0.092
(79.2, 84.2]%	0.965	-0.035
(84.2, 89.1]%	0.512	-0.488
(89.1, 94]%	0.289	-0.711
(94, 99]%	0.017	-0.983

--- Relative risk and risk increase for land coverage ---
relative_risk relative_risk_increase

Feature	relative_risk	relative_risk_increase
All lc_ are 0	0.024	-0.976
lc_dev_0	2.802	1.802
lc_dev_1	1.570	0.570
lc_dev_2	1.077	0.077
lc_dev_3	0.543	-0.457
lc_is_water/wetland	0.122	-0.878
lc_is_forest/shrub	0.923	-0.077
lc_is_grass/crop	0.270	-0.730

lc_dev_0 - Developed land, intensity level 0 (lowest, e.g. open space, very sparse development).

lc_dev_1 - Developed land, low intensity (small buildings, light residential areas).

lc_dev_2 - Developed land, medium intensity (denser housing, small commercial areas).

lc_dev_3 - Developed land, high intensity (urban cores, industrial areas, downtown).

lc_is_water/wetland - Areas covered by water (lakes, rivers) or wetlands.

lc_is_forest/shrub - Areas covered by forests or shrubland.

lc_is_grass/crop - Areas dominated by grassland or cropland (agriculture, pasture).

B.2 Risk Summary

Here, the relative risk of a power outage is based on the distribution of the history feature.

The following patterns were observed:

- Wind Gust (GUST):
- Risk increases drastically with speed.
- GUST > 28.3 → risk ×280
- GUST 17.5-20.3 → risk ×3.7
- GUST < 6.1 → risk ×0.07
- Dew Point (DPT):
- DPT 247-250 K → highest risk (×2.2)
- DPT < 240 or > 260 K → lower risk
- Precipitation Rate (PRATE):
- PRATE > 0.004 → risk ×20
- PRATE near 0 → low risk
- Population Density:
- Low density (< 600/km²) → higher outage risk (×1.9)
- High density (> 4,500/km²) → low risk (×0.3)
- Time of Day:
- Hours 15-17 (afternoon/evening) → highest risk
- Early morning hours → low risk

B.3 The Gradient Boosting Decision Tree (GBDT)-derived Rules Context

Tree #0

IF Gust ≤ 17.54m/s AND Population Density ≤ 1.02persons/km² THEN Probability of power outage is
↪ 0.117

IF Gust ≤ 17.54m/s AND Population Density > 1.02persons/km² THEN Probability of power outage is
↪ 0.109

IF Gust > 17.54m/s AND Gust ≤ 20.30m/s THEN Probability of power outage is 0.123

IF Gust > 17.54m/s AND Gust > 20.30m/s THEN Probability of power outage is 0.134

Tree #1

IF Gust ≤ 16.52m/s AND Town in {"ASHFIELD", "BECKET", "CHARLEMONT", "COLRAIN", "EGREMONT",
↪ "GRANVILLE",

"NEW MARLBOROUGH", "OTIS", "PAXTON", "PLYMPTON", "SANDISFIELD", "SAVOY", "SUDBURY", "WEST BROOKFIELD",

```

"WEST STOCKBRIDGE", "WINDSOR"} THEN Probability of power outage is 0.542
IF Gust <= 16.52m/s AND Town not in {"ASHFIELD", "BECKET", "CHARLEMONT", "COLRAIN", "EGREMONT",
↪ "GRANVILLE",
"NEW MARLBOROUGH", "OTIS", "PAXTON", "PLYMPTON", "SANDISFIELD", "SAVOY", "SUDBURY", "WEST BROOKFIELD",
"WEST STOCKBRIDGE", "WINDSOR"} THEN Probability of power outage is 0.497
IF Gust > 16.52m/s AND Gust <= 19.68m/s THEN Probability of power outage is 0.524
IF Gust > 16.52m/s AND Gust > 19.68m/s THEN Probability of power outage is 0.545
-----

```

Tree #2

```

IF Gust <= 15.53m/s AND Town in {"ASHFIELD", "BECKET", "CHARLEMONT", "CHESTERFIELD", "COLRAIN",
↪ "CONWAY",
"EGREMONT", "GRANVILLE", "HUNTINGTON", "NEW MARLBOROUGH", "OTIS", "PAXTON", "RICHMOND", "ROYALSTON",
"SANDISFIELD", "SAVOY", "SUDBURY", "WEST BROOKFIELD", "WEST STOCKBRIDGE", "WESTON", "WINDSOR"}
THEN Probability of power outage is 0.533
IF Gust <= 15.53m/s AND Town not in {"ASHFIELD", "BECKET", "CHARLEMONT", "CHESTERFIELD", "COLRAIN",
↪ "CONWAY",
"EGREMONT", "GRANVILLE", "HUNTINGTON", "NEW MARLBOROUGH", "OTIS", "PAXTON", "RICHMOND", "ROYALSTON",
"SANDISFIELD", "SAVOY", "SUDBURY", "WEST BROOKFIELD", "WEST STOCKBRIDGE", "WESTON", "WINDSOR"}
THEN Probability of power outage is 0.496
IF Gust > 15.53m/s AND month_progres_sin <= -0.68 THEN Probability of power outage is 0.542
IF Gust > 15.53m/s AND month_progres_sin > -0.68 THEN Probability of power outage is 0.521
-----

```

Tree #3

```

IF Gust <= 15.11m/s AND Town in {"ASHFIELD", "BECKET", "BLANDFORD", "CHARLEMONT", "CHESTERFIELD",
↪ "COLRAIN",
"CONWAY", "EGREMONT", "GRANVILLE", "HUNTINGTON", "LEVERETT", "LUNENBURG", "MONSON", "NEW MARLBOROUGH",
"NORTHFIELD", "OTIS", "PAXTON", "PLYMPTON", "RICHMOND", "ROCHESTER", "ROYALSTON", "SANDISFIELD",
↪ "SAVOY",
"SHEFFIELD", "STOCKBRIDGE", "SUDBURY", "WEST BROOKFIELD", "WEST STOCKBRIDGE", "WESTON", "WINDSOR"}
THEN Probability of power outage is 0.525
IF Gust <= 15.11m/s AND Town not in {"ASHFIELD", "BECKET", "BLANDFORD", "CHARLEMONT", "CHESTERFIELD",
"COLRAIN", "CONWAY", "EGREMONT", "GRANVILLE", "HUNTINGTON", "LEVERETT", "LUNENBURG", "MONSON",
"NEW MARLBOROUGH", "NORTHFIELD", "OTIS", "PAXTON", "PLYMPTON", "RICHMOND", "ROCHESTER", "ROYALSTON",
"SANDISFIELD", "SAVOY", "SHEFFIELD", "STOCKBRIDGE", "SUDBURY", "WEST BROOKFIELD", "WEST STOCKBRIDGE",
"WESTON", "WINDSOR"} THEN Probability of power outage is 0.496
IF Gust > 15.11m/s AND Gust <= 19.16m/s THEN Probability of power outage is 0.516
IF Gust > 15.11m/s AND Gust > 19.16m/s THEN Probability of power outage is 0.533
-----

```

Tree #4

```

IF Gust <= 14.41m/s AND Town in {"ASHFIELD", "CHARLEMONT", "COLRAIN", "EGREMONT", "GRANVILLE",
↪ "LEVERETT",
"NEW MARLBOROUGH", "OTIS", "PAXTON", "PLYMPTON", "SANDISFIELD", "SAVOY", "SUDBURY", "WINDSOR"}
THEN Probability of power outage is 0.529
IF Gust <= 14.41m/s AND Town not in {"ASHFIELD", "CHARLEMONT", "COLRAIN", "EGREMONT", "GRANVILLE",
↪ "LEVERETT",
"NEW MARLBOROUGH", "OTIS", "PAXTON", "PLYMPTON", "SANDISFIELD", "SAVOY", "SUDBURY", "WINDSOR"}
THEN Probability of power outage is 0.496
IF Gust > 14.41m/s AND month_progres_sin <= -0.68 THEN Probability of power outage is 0.533
IF Gust > 14.41m/s AND month_progres_sin > -0.68 THEN Probability of power outage is 0.515
-----

```

Tree #5

```

IF Gust <= 13.96m/s AND Town in {"ASHFIELD", "BECKET", "BLANDFORD", "BUCKLAND", "CHARLEMONT",
↪ "COLRAIN",
"CONWAY", "EGREMONT", "GRANVILLE", "HARDWICK", "HUNTINGTON", "LENOX", "LEVERETT", "MONTEREY",
"NEW MARLBOROUGH", "OTIS", "PAXTON", "PLYMPTON", "RICHMOND", "ROCHESTER", "SANDISFIELD", "SAVOY",
"SHEFFIELD", "SUDBURY", "WEST BROOKFIELD", "WEST STOCKBRIDGE", "WESTON", "WINDSOR"}
THEN Probability of power outage is 0.521
IF Gust <= 13.96m/s AND Town not in {"ASHFIELD", "BECKET", "BLANDFORD", "BUCKLAND", "CHARLEMONT",
↪ "COLRAIN",
"CONWAY", "EGREMONT", "GRANVILLE", "HARDWICK", "HUNTINGTON", "LENOX", "LEVERETT", "MONTEREY",
"NEW MARLBOROUGH", "OTIS", "PAXTON", "PLYMPTON", "RICHMOND", "ROCHESTER", "SANDISFIELD", "SAVOY",
"SHEFFIELD", "SUDBURY", "WEST BROOKFIELD", "WEST STOCKBRIDGE", "WESTON", "WINDSOR"}
THEN Probability of power outage is 0.496
IF Gust > 13.96m/s AND Population Density <= 1.77persons/km^2 THEN Probability of power outage is
↪ 0.529
IF Gust > 13.96m/s AND Population Density > 1.77persons/km^2 THEN Probability of power outage is 0.512
-----

```

Tree #6

```

IF 1-hour Max Wind Forecast <= 7.57m/s AND Population Density <= 1.02persons/km^2
THEN Probability of power outage is 0.512

```

IF 1-hour Max Wind Forecast <= 7.57m/s AND Population Density > 1.02persons/km^2
THEN Probability of power outage is 0.496
IF 1-hour Max Wind Forecast > 7.57m/s AND lunary_progress_sin <= 0.89
THEN Probability of power outage is 0.514
IF 1-hour Max Wind Forecast > 7.57m/s AND lunary_progress_sin > 0.89
THEN Probability of power outage is 0.528

Tree #7

IF 1-hour Max Wind Forecast <= 8.00m/s AND 1-hour Max Wind Forecast <= 5.74m/s
THEN Probability of power outage is 0.497
IF 1-hour Max Wind Forecast <= 8.00m/s AND 1-hour Max Wind Forecast > 5.74m/s
THEN Probability of power outage is 0.504
IF 1-hour Max Wind Forecast > 8.00m/s AND lunary_progress_sin <= 0.87
THEN Probability of power outage is 0.515
IF 1-hour Max Wind Forecast > 8.00m/s AND lunary_progress_sin > 0.87
THEN Probability of power outage is 0.526

Tree #8

IF Gust <= 13.82m/s AND Population Density <= 1.02persons/km^2 THEN Probability of power outage is
↪ 0.510
IF Gust <= 13.82m/s AND Population Density > 1.02persons/km^2 THEN Probability of power outage is
↪ 0.496
IF Gust > 13.82m/s AND Population Density <= 5.08persons/km^2 THEN Probability of power outage is
↪ 0.521
IF Gust > 13.82m/s AND Population Density > 5.08persons/km^2 THEN Probability of power outage is 0.507

Tree #9

IF Gust <= 13.82m/s AND Town in {"ASHFIELD", "BECKET", "BLANDFORD", "CHARLEMONT", "CHESTERFIELD",
↪ "COLRAIN",
"CONWAY", "DOVER", "EGREMONT", "GRANVILLE", "GREAT BARRINGTON", "HANCOCK", "HARDWICK", "HARVARD",
"LEVERETT", "LINCOLN", "LUNENBURG", "NEW MARLBOROUGH", "OTIS", "PAXTON", "RICHMOND", "ROCHESTER",
"SANDISFIELD", "SAVOY", "SHEFFIELD", "STOCKBRIDGE", "SUDBURY", "WEST BROOKFIELD", "WEST STOCKBRIDGE",
"WESTON", "WINDSOR"}
THEN Probability of power outage is 0.517
IF Gust <= 13.82m/s AND Town not in {"ASHFIELD", "BECKET", "BLANDFORD", "CHARLEMONT", "CHESTERFIELD",
"COLRAIN", "CONWAY", "DOVER", "EGREMONT", "GRANVILLE", "GREAT BARRINGTON", "HANCOCK", "HARDWICK",
↪ "HARVARD",
"LEVERETT", "LINCOLN", "LUNENBURG", "NEW MARLBOROUGH", "OTIS", "PAXTON", "RICHMOND", "ROCHESTER",
"SANDISFIELD", "SAVOY", "SHEFFIELD", "STOCKBRIDGE", "SUDBURY", "WEST BROOKFIELD",
"WEST STOCKBRIDGE", "WESTON", "WINDSOR"}
THEN Probability of power outage is 0.496
IF Gust > 13.82m/s AND month_progres_sin <= -0.68 THEN Probability of power outage is 0.523
IF Gust > 13.82m/s AND month_progres_sin > -0.68 THEN Probability of power outage is 0.509

Tree #10

IF 1-hour Max Wind Forecast <= 6.94m/s AND Town in {"ASHFIELD", "BECKET", "BLANDFORD", "CHARLEMONT",
"CHESTERFIELD", "COLRAIN", "CONWAY", "DEERFIELD", "EGREMONT", "GRANVILLE", "GREAT BARRINGTON",
↪ "HARDWICK",
"HUNTINGTON", "LEVERETT", "LUNENBURG", "MILLVILLE", "NEW MARLBOROUGH", "NORTHFIELD", "OTIS", "PAXTON",
"PLYMPTON", "ROCHESTER", "ROYALSTON", "SANDISFIELD", "SAVOY", "SHUTESBURY", "SUDBURY", "WAYLAND",
"WEST BROOKFIELD", "WEST STOCKBRIDGE", "WESTON", "WINDSOR"} THEN Probability of power outage is 0.517
IF 1-hour Max Wind Forecast <= 6.94m/s AND Town not in {"ASHFIELD", "BECKET", "BLANDFORD",
↪ "CHARLEMONT",
"CHESTERFIELD", "COLRAIN", "CONWAY", "DEERFIELD", "EGREMONT", "GRANVILLE", "GREAT BARRINGTON",
↪ "HARDWICK",
"HUNTINGTON", "LEVERETT", "LUNENBURG", "MILLVILLE", "NEW MARLBOROUGH", "NORTHFIELD", "OTIS", "PAXTON",
"PLYMPTON", "ROCHESTER", "ROYALSTON", "SANDISFIELD", "SAVOY", "SHUTESBURY", "SUDBURY", "WAYLAND",
"WEST BROOKFIELD", "WEST STOCKBRIDGE", "WESTON", "WINDSOR"} THEN Probability of power outage is 0.496
IF 1-hour Max Wind Forecast > 6.94m/s AND Gust <= 18.27m/s THEN Probability of power outage is 0.508
IF 1-hour Max Wind Forecast > 6.94m/s AND Gust > 18.27m/s THEN Probability of power outage is 0.520

Tree #11

IF Gust <= 13.82m/s AND Town in {"ASHFIELD", "BECKET", "BLANDFORD", "CHARLEMONT", "CHARLTON",
↪ "CHESTERFIELD",
"COLRAIN", "CONWAY", "EGREMONT", "GRANVILLE", "HANCOCK", "HARDWICK", "HARVARD", "HUNTINGTON", "LENOX",
"LEVERETT", "LUNENBURG", "MONSON", "MONTEREY", "NEW MARLBOROUGH", "OTIS", "PAXTON", "PLYMPTON",
↪ "RICHMOND",
"SANDISFIELD", "SAVOY", "STOCKBRIDGE", "SUDBURY", "WAYLAND", "WEST BROOKFIELD", "WESTON", "WINDSOR"}
THEN Probability of power outage is 0.515
IF Gust <= 13.82m/s AND Town not in {"ASHFIELD", "BECKET", "BLANDFORD", "CHARLEMONT", "CHARLTON",

```

"CHESTERFIELD", "COLRAIN", "CONWAY", "EGREMONT", "GRANVILLE", "HANCOCK", "HARDWICK", "HARVARD",
"HUNTINGTON", "LENOX", "LEVERETT", "LUNENBURG", "MONSON", "MONTEREY", "NEW MARLBOROUGH", "OTIS",
↪ "PAXTON",
"PLYMPTON", "RICHMOND", "SANDISFIELD", "SAVOY", "STOCKBRIDGE", "SUDBURY", "WAYLAND", "WEST
↪ BROOKFIELD",
"WESTON", "WINDSOR"}
THEN Probability of power outage is 0.496
IF Gust > 13.82m/s AND Population Density <= 5.08persons/km^2 THEN Probability of power outage is
↪ 0.517
IF Gust > 13.82m/s AND Population Density > 5.08persons/km^2 THEN Probability of power outage is 0.506
-----
Tree #12
IF 1-hour Max Wind Forecast <= 6.63m/s AND Town in {"ASHFIELD", "BECKET", "BLANDFORD", "BUCKLAND",
"CHARLEMONT", "CHESTERFIELD", "COLRAIN", "CONWAY", "EGREMONT", "GRANVILLE", "GREAT BARRINGTON",
"HANCOCK", "HUNTINGTON", "LEVERETT", "MILLVILLE", "OTIS", "PAXTON", "RICHMOND", "ROCHESTER",
"ROYALSTON", "SANDISFIELD", "SAVOY", "SHEFFIELD", "SHUTESBURY", "STOCKBRIDGE", "SUDBURY", "WEST
↪ BROOKFIELD",
"WEST STOCKBRIDGE", "WESTON", "WILBRAHAM", "WINDSOR"} THEN Probability of power outage is 0.515
IF 1-hour Max Wind Forecast <= 6.63m/s AND Town not in {"ASHFIELD", "BECKET", "BLANDFORD", "BUCKLAND",
"CHARLEMONT", "CHESTERFIELD", "COLRAIN", "CONWAY", "EGREMONT", "GRANVILLE", "GREAT BARRINGTON",
↪ "HANCOCK",
"HUNTINGTON", "LEVERETT", "MILLVILLE", "OTIS", "PAXTON", "RICHMOND", "ROCHESTER", "ROYALSTON",
↪ "SANDISFIELD",
"SAVOY", "SHEFFIELD", "SHUTESBURY", "STOCKBRIDGE", "SUDBURY", "WEST BROOKFIELD", "WEST STOCKBRIDGE",
"WESTON", "WILBRAHAM", "WINDSOR"}
THEN Probability of power outage is 0.496
IF 1-hour Max Wind Forecast > 6.63m/s AND Population Density <= 10.60persons/km^2
THEN Probability of power outage is 0.515
IF 1-hour Max Wind Forecast > 6.63m/s AND Population Density > 10.60persons/km^2
THEN Probability of power outage is 0.503
-----
Tree #13
IF Gust <= 13.82m/s AND Population Density <= 1.77persons/km^2 THEN Probability of power outage is
↪ 0.506
IF Gust <= 13.82m/s AND Population Density > 1.77persons/km^2 THEN Probability of power outage is
↪ 0.496
IF Gust > 13.82m/s AND month_progres_sin <= -0.68 THEN Probability of power outage is 0.519
IF Gust > 13.82m/s AND month_progres_sin > -0.68 THEN Probability of power outage is 0.508
-----
Tree #14
IF 1-hour Max Wind Forecast <= 6.94m/s AND Town in {"ASHFIELD", "BECKET", "BLANDFORD", "BROOKFIELD",
"BUCKLAND", "CHARLEMONT", "CHESTERFIELD", "COLRAIN", "CONWAY", "DEERFIELD", "EGREMONT", "GRANVILLE",
"HANCOCK", "HUBBARDSTON", "LANESBOROUGH", "LENOX", "LEVERETT", "LUNENBURG", "MONSON", "NEW
↪ MARLBOROUGH",
"NORTHFIELD", "PAXTON", "RICHMOND", "ROYALSTON", "SANDISFIELD", "SAVOY", "SHEFFIELD", "SHUTESBURY",
"SUDBURY", "WEST BROOKFIELD", "WESTON", "WINDSOR"}
THEN Probability of power outage is 0.514
IF 1-hour Max Wind Forecast <= 6.94m/s AND Town not in {"ASHFIELD", "BECKET", "BLANDFORD",
↪ "BROOKFIELD",
"BUCKLAND", "CHARLEMONT", "CHESTERFIELD", "COLRAIN", "CONWAY", "DEERFIELD", "EGREMONT", "GRANVILLE",
"HANCOCK", "HUBBARDSTON", "LANESBOROUGH", "LENOX", "LEVERETT", "LUNENBURG", "MONSON", "NEW
↪ MARLBOROUGH",
"NORTHFIELD", "PAXTON", "RICHMOND", "ROYALSTON", "SANDISFIELD", "SAVOY", "SHEFFIELD", "SHUTESBURY",
"SUDBURY", "WEST BROOKFIELD", "WESTON", "WINDSOR"}
THEN Probability of power outage is 0.497
IF 1-hour Max Wind Forecast > 6.94m/s AND Gust <= 19.68m/s THEN Probability of power outage is 0.507
IF 1-hour Max Wind Forecast > 6.94m/s AND Gust > 19.68m/s THEN Probability of power outage is 0.518
-----

```

C Qualitative Analysis

Table 5 illustrates how different context influence the model's risk assessments. **No context** provides only a basic interpretation using the models' internal parametric knowledge, leading to a moderate risk estimate based on the weather conditions. **Quantitative Risk Context** uses structured numerical features, resulting in a high-risk prediction driven by extreme wind speed and precipitation. **Risk Summary Context** also integrate textual information, also producing a high-risk assessment. **GBDT-derived Rule** concludes a low risk because it only uses part of decision trees, which can not fully capture the complexity of the complete model.

Query: Today is January 23, Monday, and it is not a holiday. It's cold. The air is dry. There are strong wind gusts. It is raining. There has been recent snowfall. There is some wind. The location is in QUINCY, Massachusetts. The area has a moderate population density. The utility group is NationalGrid. The location is in the middle of a city, with tall buildings everywhere, mostly concrete, and roads with little open land. There is very little tree canopy coverage.

Context	RESPONSE
NO CONTEXT	{“risk_level”: “medium”, “Explanation”: “Strong wind gusts and rain could cause tree branches to fall, potentially damaging power lines in an area with little tree canopy coverage.”}
+ QUANTITATIVE RISK CONTEXT	<p>{“risk_level”: “high”, “Explanation”: “Strong wind gusts, cold temperatures, dry air, and heavy precipitation create a high risk of power outages.”} Reasoning:</p> <ul style="list-style-type: none"> • Wind Gusts: Likely above 28.3 m/s — may damage power lines. • Cold Temperatures: Risk of ice on lines; system stress. • Dry Air: Exacerbates wind/cold effects. • Heavy Precipitation: May cause short circuits. • Urban Environment: Dense infrastructure increases impact. • Low Tree Canopy: Less insulation from wind/weather. <p>These combined factors contribute to a high risk.</p>
+ RISK SUMMARY CONTEXT	<p>{“risk_level”: “high”, “Explanation”: “Strong wind gusts and precipitation increase the risk significantly.”} Reasoning:</p> <ul style="list-style-type: none"> • Wind Gusts (GUST): The description mentions "strong wind gusts," which puts the GUST value in the range where the risk increases drastically (GUST > 28.3 → risk ×280). • Precipitation Rate (PRATE): It is raining (PRATE > 0.004 → risk ×20), further increasing the risk. • Population Density: The area has a "moderate" population density, which does not significantly alter the risk level. • Time of Day: Not specified, but being in the middle of a city means moderate to high risk at any time due to the urban environment. • Environmental Conditions: Cold and dry air do not significantly impact the risk based on given patterns. • Recent Snowfall: While recent snowfall may affect road conditions, it is not directly included as a factor in the risk calculation. <p>Given the significant contributions from strong wind gusts and precipitation, the overall risk level is classified as “high.”</p>
+ GBDT-DERIVED RULE CONTEXT	<p>{“risk_level”: “low”, “reason”: “The location in Quincy, Massachusetts, with its moderate population density and urban environment, reduces the likelihood of power outages despite strong winds.”} Reasoning:</p> <ul style="list-style-type: none"> • Location: Quincy is an urban area with tall buildings and little open land, which generally means fewer trees that could fall and cause power outages. • Population Density: Moderate population density suggests a more developed area where infrastructure is likely better maintained. • Wind Conditions: While there are strong wind gusts, the urban environment provides natural protection against significant damage. • Other Factors: The presence of tall buildings and concrete structures can help shield power lines from direct impact, reducing the risk of outages. <p>Given these factors, the overall risk level is considered “low.”</p>

Table 5: Qualitative Analysis of Qwen2.5-7B Across Different Contexts.