1 Further Details about the Motivation of KBMC

Estraction of word-level probabilistic significance is a powerful in identifying frequent and significant patterns within large datasets [5, 9], and its application to fake news detection introduces several compelling advantages, particularly in the realm of uncertainty quantification [6, 2]. The use of word-level probabilistic significance-based dropout in uncertainty quantification for fake news detection can significantly enhance the robustness and reliability of fake news detection systems. The rapid proliferation of fake news on digital platforms poses a significant challenge to information integrity and societal trust. Traditional approaches to fake news detection, including machine learning [4, 1, 10] and deep learning techniques [8, 13, 7], focus on classifying news articles as true or false based on their textual content, structure, and dissemination patterns. However, these methods often struggle with uncertainty in classification, particularly in ambiguous or borderline cases where distinguishing between true and false information is inherently difficult. Word-level probabilistic significance extraction identifies frequent and significant patterns in large datasets, enabling the discovery of common features and trends associated with fake and real news [3, 12]. By systematically analyzing these patterns, we can enhance the robustness of detection models. For instance, recurrent phrases, linguistic styles, and propagation characteristics of fake news can be detected and used to train more accurate and resilient models. A significant challenge in fake news detection is the inherent uncertainty in classification. By extracting probabilities from news datasets, we can quantify the uncertainty associated with each classification decision. For example, if a news article contains several frequent patterns associated with fake news, the model can flag it with higher confidence as potentially fake. In contrast, the absence of such patterns can indicate lower certainty. Monte Carlo method combined with word-level probabilistic significance-based dropout offer a robust approach for uncertainty quantification in fake news detection, enabling more reliable assessments of content veracity. By simulating numerous scenarios and extracting prevalent patterns, we can effectively handle the inherent unpredictability of information sources, leading to more accurate and confident of the fake news detection process.

2 Datasets, Experimental Settings and Workflow

We evaluate the designed method using the FakeNewsNet [11] and ISOT Fake News datasets. FakeNewsNet is a comprehensive dataset specifically created for the study of fake news detection. It comprises two distinct collections of articles sourced from two different fact-checking websites: GossipCop and PolitiFact. PolitiFact, operated by the Tampa Bay Times, addresses political statements and claims, while GossipCop, which has now been discontinued, focused on debunking entertainment and celebrity rumors. The dataset includes a total of 23,190 articles, each labeled as either real or fake. The distribution of these labels is imbalanced, with a ratio of 3:1, meaning there are three times as many real articles as fake ones. This imbalance presents a realistic challenge, as it mirrors the common real-world scenario where fake news is typically less prevalent than real news. The ISOT Fake News Dataset contains articles mostly from 2016 and 2017. Articles marked as "Real" (21417 of them) are from reuters.com (operated by Thomson Reuters) and the other 23482 come from different sources and were flagged as fake by Politifact and Wikipedia. It is worth noting that the proportion between these two sets is equal and that mistakes (e.g. spelling mistakes) were kept in the text.

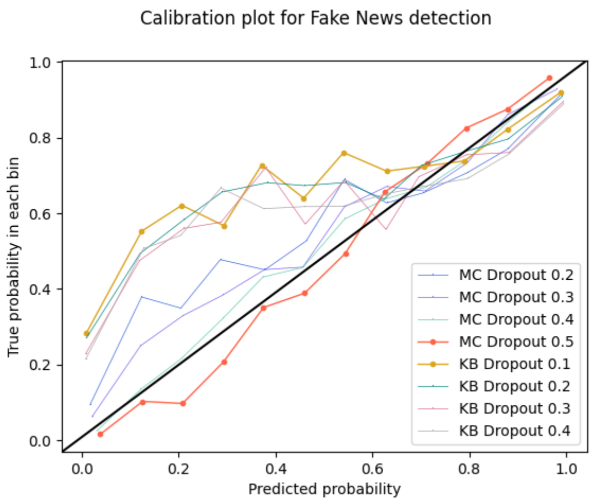

Figure 1. Calibration curves for Monte Carlo Dropout and Knowledge Base Dropout with FNN Dataset

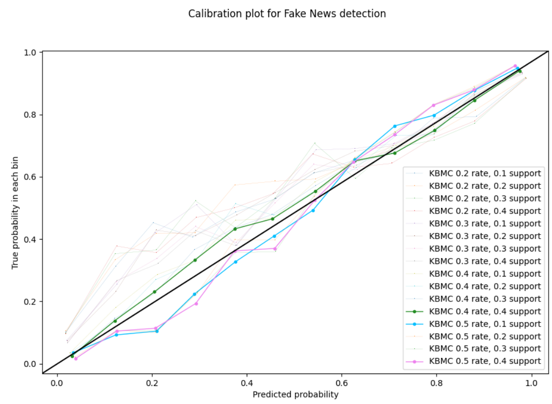

Figure 2. Calibration curves for the ablation study with FNN Dataset

An important question arises and that is whether these methods too computationally expensive. Firstly, the whole process is way longer when we use the ISOT dataset, which is due to it having much longer articles. Otherwise the results are fairly consistent, with MC Dropout, KBMC 0.4 0.4 and bootstrapping having the shortest execution times.

3 Example of Knowledge Base Creation

Consider a dataset $\mathcal{D}$ with the following news articles represented by keywords:

- T_1: [election, fraud, social media, claim]

Table 1. Comparison of execution time (in miliseconds) of the chosen $\mathcal{UQ}$ methods for both datasets

Dataset	$KB_{0.1}$	$MC_{0.5}$	$KB_{0.4}MC_{0.4}$	$KB_{0.1}MC_{0.5}$	$Bp_{0.5}$	En	$Bp_{0.5}KB_{0.1}$	$KB_{0.1}En$	$KB_{0.2}En$
FakeNewsNet	1091,619	731,814	720,333	1005,603	739,770	-	985,706	-	-
ISOT	2694,679	2111,435	2170,936	2455,887	2162,841	-	2483,342	-	-

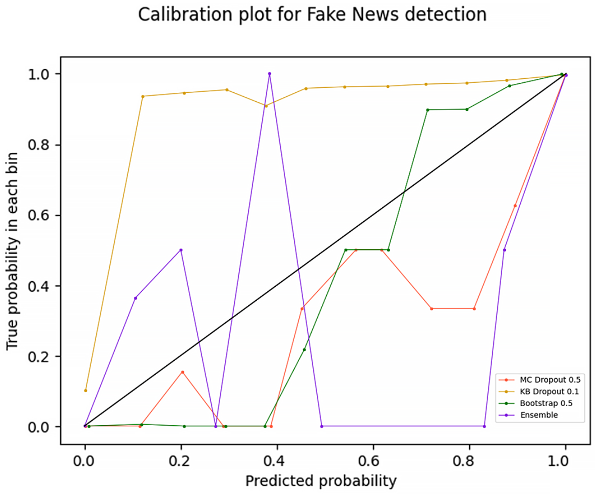

Figure 3. Calibration for base methods with ISOT Dataset

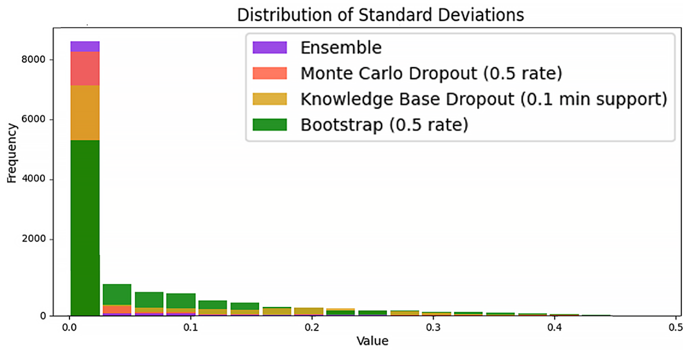

Figure 4. Standard deviations of predictions for base methods with ISOT (bars overlap)

- T_2: [election, results, certified]
- T_3: [pandemic, vaccine, health, claim]
- T_4: [election, fraud, recount]
- T_5: [pandemic, health, misinformation]

Assume a minimum support threshold min_sup = 0.4 (40%).

Step 1: Generate Frequent 1-Itemsets Calculate the support for each individual keyword:

- $\sigma(\text{election}) = \frac{3}{5} = 0.6$
- $\sigma(\text{fraud}) = \frac{2}{5} = 0.4$
- $\sigma(\text{social media}) = \frac{1}{5} = 0.2$
- $\sigma(\text{claim}) = \frac{2}{5} = 0.4$
- $\sigma(\text{results}) = \frac{1}{5} = 0.2$
- $\sigma(\text{certified}) = \frac{1}{5} = 0.2$
- $\sigma(\text{pandemic}) = \frac{2}{5} = 0.4$
- $\sigma(\text{vaccine}) = \frac{1}{5} = 0.2$
- $\sigma(\text{health}) = \frac{2}{5} = 0.4$
- $\sigma(\text{recount}) = \frac{1}{5} = 0.2$

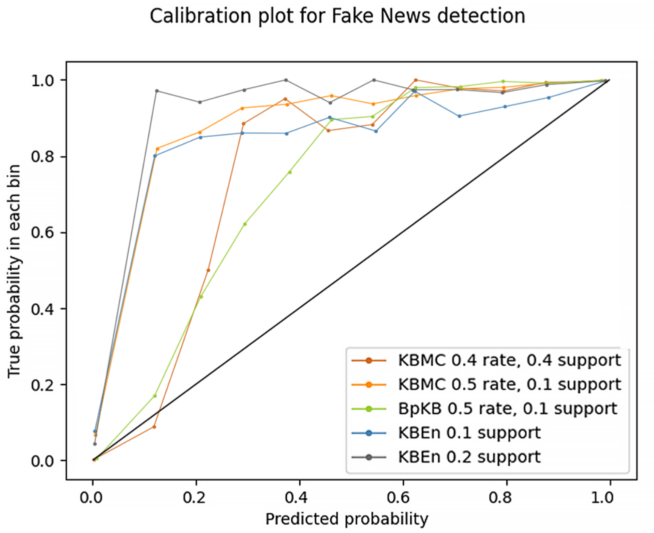

Figure 5. Calibration for combined methods with ISOT Dataset

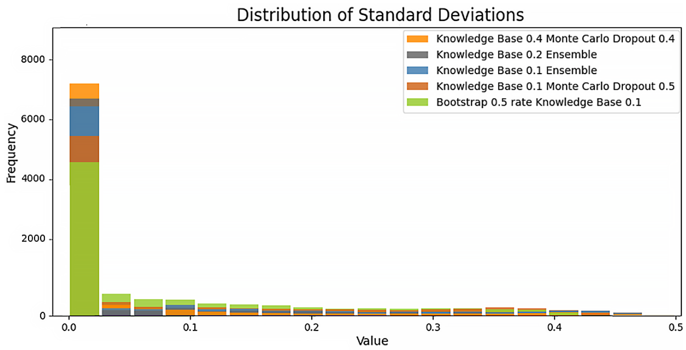

Figure 6. Standard deviations of predictions for combined methods with ISOT (bars overlap)

- $\sigma(\text{misinformation}) = \frac{1}{5} = 0.2$

Frequent 1-itemsets with support ≥ 0.4 is: {election, fraud, claim, pandemic, health}.

Step 2: Generate Frequent 2-Itemsets Generate candidate 2-itemsets from frequent 1-itemsets and calculate their supports:

- $\sigma(\{\text{election, fraud}\}) = \frac{2}{5} = 0.4$
- $\sigma(\{\text{election, claim}\}) = \frac{1}{5} = 0.2$
- $\sigma(\{\text{election, pandemic}\}) = \frac{0}{5} = 0.0$
- $\sigma(\{\text{election, health}\}) = \frac{0}{5} = 0.0$
- $\sigma(\{\text{fraud, claim}\}) = \frac{1}{5} = 0.2$
- $\sigma(\{\text{fraud, pandemic}\}) = \frac{0}{5} = 0.0$
- $\sigma(\{\text{fraud, health}\}) = \frac{0}{5} = 0.0$
- $\sigma(\{\text{claim, pandemic}\}) = \frac{1}{5} = 0.2$
- $\sigma(\{\text{claim, health}\}) = \frac{1}{5} = 0.2$
- $\sigma(\{\text{pandemic, health}\}) = \frac{2}{5} = 0.4$

Table 2. Patterns from real and fake news

	Pattern	Support
Real news	σ(one)	0.479215641304971
	σ(time)	0.47520210997075857
	σ(also)	0.4391376641247635
	σ(like)	0.43013588670374403
	σ(year)	0.4295625250845708
	σ(show)	0.41666188865317355
	σ(said)	0.4156871739005791
	σ(new)	0.40995355770884695
Fake news	σ(time)	0.48618592528236315
	σ(one)	0.4536924413553432
	σ(like)	0.4300608166811468
	σ(year)	0.4297132927888792
	σ(get)	0.4229365768896612
	σ(said)	0.4161598609904431
	σ(new)	0.4005212858384014

Frequent 2-itemsets with support $\geq$ 0.4 is: { {election, fraud}, {pandemic, health} }.

Step 3: Generate Frequent 3-Itemsets Generate candidate 3-itemsets from frequent 2-itemsets and calculate their supports:

- $\sigma(\{\text{election, fraud, claim}\}) = \frac{1}{5} = 0.2$
- $\sigma(\{\text{pandemic, health, claim}\}) = \frac{1}{5} = 0.2$

No frequent 3-itemsets with support $\geq$ 0.4.

The extracted frequent patterns can be integrated into a knowledge base to enhance defect detection. For instance, if the pattern {election, fraud} is frequent in fake news, the knowledge base can use this information to flag news articles containing these terms as potentially fake. Similarly, patterns like {pandemic, health} may be used to verify the authenticity of health-related news.

After running the model multiple times, the average of results for each article is calculated.

4 Example of Uncertainty Quantification using KBMC

Consider a following article:

"Ronald claims that the recent elections were unjust, alleging instances of electoral fraud, and is calling for a recount via social media."

We match the preprocessed set of words from this article with the previously mentioned dataset $\mathcal{D}$, containing the support for each individual keyword:

- $\sigma(\text{election}) = \frac{3}{5} = 0.6$
- $\sigma(\text{fraud}) = \frac{2}{5} = 0.4$
- $\sigma(\text{social media}) = \frac{1}{5} = 0.2$
- $\sigma(\text{claim}) = \frac{2}{5} = 0.4$
- $\sigma(\text{recount}) = \frac{1}{5} = 0.2$

In our method, each of these words have a chance to be dropped out in a run of the model with probability equal to its support score. So in this example, the word "election" has a 0.6% chance to be dropped out, while "recount" has a 0.2% chance. MC Dropout and KB Dropout have been compared as both seperate and joint techniques for uncertainty quantification. As seen in Figure **??**, KB Dropout is overly uncertain for values 0.0 - 0.6 of predicted probability, but both MC and KB's performance increases when reaching higher values. We have run the models using different values of dropout rate and minimum support in order to identify those, which have the best results. Figure additionally **??** illustrates the performance of each version of the model and we can see that the chosen KBMC models are the closest to the perfect performance, i.e., have the lowest calibration errors.

Qualitative Results of KBMC The Apriori algorithm is meant to extract most relevant patterns in sets of data. Listed below are most frequent patterns that appear in fake and real news, with minimum support of 0.4. We can hypothesize that these patterns are actually characteristic for news reporting in general and that removing them from the input increases the importance of actual differences between real and fake news. To test this, however, we would need to extract patterns with statistically insignificantly different support in fake and real news and see how it affects the results.

References

[1] A. M. Braşoveanu and R. Andonie. Integrating machine learning techniques in semantic fake news detection. *Neural Processing Letters*, 53 (5):3055–3072, 2021.

[2] W. Chen, B. Zhang, and M. Lu. Uncertainty quantification for multilabel text classification. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 10(6):e1384, 2020.

[3] Y. Dong, D. He, X. Wang, Y. Jin, M. Ge, C. Yang, and D. Jin. Unveiling implicit deceptive patterns in multi-modal fake news via neuro-symbolic reasoning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pages 8354–8362, 2024.

[4] Y. Dou, K. Shu, C. Xia, P. S. Yu, and L. Sun. User preference-aware fake news detection. In *Proceedings of the 44th international ACM SIGIR conference on research and development in information retrieval*, pages 2051–2055, 2021.

[5] G. Huang, W. Gan, and P. S. Yu. Taspm: Targeted sequential pattern mining. *ACM Transactions on Knowledge Discovery from Data*, 18(5): 1–18, 2024.

[6] C. Kamath. On the role of data mining techniques in uncertainty quantification. *International Journal for Uncertainty Quantification*, 2(1), 2012.

[7] Q. Liu, J. Wu, S. Wu, and L. Wang. Out-of-distribution evidence-aware fake news detection via dual adversarial debiasing. *IEEE Transactions on Knowledge and Data Engineering*, 2024.

[8] A. Mosallanezhad, M. Karami, K. Shu, M. V. Mancenido, and H. Liu. Domain adaptive fake news detection via reinforcement learning. In *Proceedings of the ACM Web Conference 2022*, pages 3632–3640, 2022.

[9] R. Pérez-Chacón, G. Asencio-Cortés, A. Troncoso, and F. Martínez-Álvarez. Pattern sequence-based algorithm for multivariate big data time series forecasting: Application to electricity consumption. *Future Generation Computer Systems*, 154:397–412, 2024.

[10] Q. Sheng, X. Zhang, J. Cao, and L. Zhong. Integrating pattern-and fact-based fake news detection via model preference learning. In *Proceedings of the 30th ACM international conference on information & knowledge management*, pages 1640–1650, 2021.

[11] K. Shu, D. Mahudeswaran, S. Wang, D. Lee, and H. Liu. Fakenewsnet: A data repository with news content, social context, and spatiotemporal information for studying fake news on social media. *Big data*, 8(3): 171–188, 2020.

[12] X. Su, J. Yang, J. Wu, and Y. Zhang. Mining user-aware multi-relations for fake news detection in large scale online social networks. In *Proceedings of the sixteenth ACM international conference on web search and data mining*, pages 51–59, 2023.

[13] L. Xiao, Q. Zhang, C. Shi, S. Wang, U. Naseem, and L. Hu. Msynfd: Multi-hop syntax aware fake news detection. In *Proceedings of the ACM on Web Conference 2024*, pages 4128–4137, 2024.