Received: 8 January 2026 | Revised: 27 April 2026 | Accepted: 18 May 2026 | Published online: 24 July 2026

RESEARCH ARTICLE

Journal of Computational and Cognitive Engineering
2026, Vol. 00(00) 1–12
DOI: 10.47852/bonviewJCCE62029040

BON VIEW PUBLISHING

A Hybrid Transformer–Ontology Framework for Halal Food Classification Using Multilingual Ingredient Label Analysis

Mohd Azmi Al Betar[1], Noorrezam Yusop[2,*], Tao Hai[1], Nor Aiza Moketar[2], Nuridawati Mustafa[2], Mohd Nazrien Zaraini[3] and Massila Kamalrudin[2]

[1] *Artificial Intelligence Research Centre, Ajman University, United Arab Emirates*

[2] *Department of Software Engineering, Universiti Teknikal Malaysia Melaka, Malaysia*

[3] *Department of Computer System and Communication, Universiti Teknikal Malaysia Melaka, Malaysia*

Abstract: Ensuring halal compliance within global food supply chains has become increasingly challenging due to multilingual ingredient labeling, ambiguous additive terminology, and heterogeneous sourcing practices. Existing rule-based systems or sequential machine learning models have limited capacity to resolve semantic ambiguity, address cross-lingual variation, and meet the interpretability requirements imposed by halal certification authorities. A hybrid neural–symbolic classification framework is proposed that integrates a multilingual transformer-based semantic encoder (XLM-R) with ontology-driven lexical reasoning for automated halal ingredient assessment. The proposed architecture leverages contextual embeddings learned from Malay, Arabic, and English ingredient descriptions, while a structured halal ontology encodes domain knowledge on ingredient origin, permissibility, and confidence levels. A probabilistic fusion mechanism combines transformer prediction scores with weighted ontological evidence to produce confidence-aware classifications into halal, haram, and syubhah categories. An experimental evaluation was conducted on a dataset of 10,000 product records using an 80:20 stratified train-test split. The proposed XLM-R + Ontology model achieves consistently superior performance, achieving overall accuracy of 0.81–1.00, recall of 0.90–1.00, and macro-averaged F1 of 0.85–0.90, outperforming Naive Bayes, support vector machine, and LSTM baselines. The hybrid framework demonstrates substantial improvements in detecting ambiguous syubhah cases and low-resource haram indicators, which remain problematic for conventional classifiers. Attention-based interpretability and performance attribution analyses indicate that integrating contextual semantic learning with structured symbolic knowledge enhances both reliability and transparency. These findings underscore the effectiveness of hybrid neural-symbolic architectures for multilingual, domain-specific food classification and support deployment in consumer-oriented digital platforms.

Keywords: multilingual transformer, halal food classification, domain adaptation, semantic reasoning, halal ingredient ontology

1. Introduction

Halal dietary compliance remains a mandatory religious requirement for Muslim consumers worldwide and constitutes a rapidly expanding segment of the global food industry. This growth is driven by increasing Muslim populations and rising consumer awareness of ethically sourced and religiously compliant products [1]. Nevertheless, modern halal verification is challenged by cross-border manufacturing practices, fragmented regulatory environments, and heterogeneous food labeling standards. These factors exacerbate information asymmetry and complicate product traceability across international supply chains, thereby undermining reliable halal assurance mechanisms [2, 3].

One of the most critical difficulties in halal food verification arises from ingredient disclosure practices [4]. Food labels frequently employ complex chemical terminology, region-specific nomenclature, and additives derived from heterogeneous biological or synthetic sources, each of which may carry different halal implications. Manual verification under such conditions is labor-intensive and susceptible to classification errors, particularly in multilingual retail contexts where ingredient lists may be presented in multiple languages or technical scientific forms [5]. The absence of standardized and interoperable ingredient ontologies further intensifies semantic ambiguity, reflecting challenges similar to those encountered in pharmaceutical and biomedical terminology harmonization.

Existing computational approaches for halal classification have predominantly relied on rule-based expert systems or conventional statistical machine learning techniques [4]. Rule-based frameworks offer transparency but require extensive manual

***Corresponding author:** Noorrezam Yusop, Department of Software Engineering, Universiti Teknikal Malaysia Melaka, Malaysia. Email: noorrezam@utem.edu.my

curation and suffer from limited scalability and adaptability when addressing emerging additives, evolving manufacturing processes, and regional variations in ingredient naming conventions [1]. While traditional machine learning models demonstrate improved predictive accuracy over static rules [6], they remain constrained by shallow textual representations and limited capacity to capture contextual semantics, particularly in cases involving polysemous ingredient descriptions or multilingual transliterations.

Recent advancements in transformer-based deep learning architectures provide a promising avenue for overcoming these limitations. Transformer models utilize self-attention mechanisms to model long-range contextual dependencies and capture deep semantic relationships within text, outperforming earlier natural language processing (NLP) techniques across a range of tasks [7]. Multilingual pre-trained models such as mBERT and XLM-R further extend these capabilities by enabling cross-lingual semantic alignment through shared embedding spaces learned from large multilingual corpora. These properties are especially relevant for halal ingredient analysis, which requires accurate semantic interpretation across languages, regulatory frameworks, and culturally specific naming practices.

To address these interrelated challenges, this study proposes a hybrid halal ingredient classification framework that integrates multilingual transformer-based contextual embeddings with ontology-driven semantic reasoning. The framework incorporates systematic data preprocessing, multilingual lexical mapping, supervised transformer fine-tuning, and confidence-aware inference to classify ingredients into halal, haram, and syubhah categories [8]. By combining data-driven learning with structured domain knowledge, the proposed approach improves classification robustness, interpretability, and scalability. Compared to existing methods, the framework offers enhanced semantic understanding through transformer architectures [9], native multilingual capability via cross-lingual transfer learning using models such as XLM-R and mBART [10], and practical applicability through confidence-based prediction and iterative refinement, supporting deployment in automated halal certification systems, regulatory platforms, and consumer-oriented applications [11].

2. Literature Review

The application of artificial intelligence and NLP to halal food verification has gained increasing research attention due to the complexity of multilingual ingredient labeling and fragmented global supply chains [12]. Existing studies have explored automated systems for determining halal status through ingredient analysis, product labeling, and certification databases [13]. However, semantic ambiguity, multilingual variability, and interpretability remain unresolved challenges.

Early halal classification systems primarily employed classical machine learning techniques such as support vector machines (SVMs) trained on labeled ingredient datasets [14]. While effective at a basic level, these approaches rely on shallow textual features and exhibit limited semantic generalization, particularly in multilingual settings. Deep learning methods, including convolutional neural networks applied to food packaging images, have also been proposed [15], yet visual analysis alone cannot fully capture ingredient-level semantic compliance.

Recent studies consistently identify two critical limitations: the scarcity of standardized multilingual datasets and the lack of semantic embedding mechanisms that can resolve cross-lingual ambiguity [16]. These limitations motivate the adoption of context-aware and cross-lingual modeling techniques [16]. Empirical findings further demonstrate that multilingual semantic embedding is essential for robust halal ingredients across Malay, Arabic, and English texts [17].

Transformer-based architectures have significantly advanced NLP by enabling contextual representation learning through self-attention mechanisms [18]. Models such as BERT, RoBERTa, and DistilBERT achieve state-of-the-art performance across text classification tasks [19, 20]. In multilingual contexts, transformer models learn shared semantic spaces that align equivalent concepts across languages [21–23]. Comparative evaluations show that transformer encoders consistently outperform classical classifiers such as SVM and Naïve Bayes (NB) [24–26], with multilingual models such as XLM-R demonstrating strong performance in low-resource settings [27, 28].

Nevertheless, recent research emphasizes that transformer models alone may be insufficient for domain-sensitive tasks requiring regulatory interpretability [29], such as halal classification. Hybrid approaches integrating multilingual embeddings with bilingual lexical resources and halal ontologies have been shown to improve semantic consistency, transparency, and classification accuracy. These systems support standardized ingredient mapping, origin-based classification, and rule-aligned reasoning, which are essential for halal certification.

Excellent advances in machine learning have demonstrated the powerful ability to model complex and nonlinear relationships in a wide range of areas, driving the use of smart hybrid systems. In particular, neural network approaches have shown remarkable success in learning high-dimensional representations and in extracting semantic patterns from non-structured data [30], but, despite their powerful predictive power, these models often lack the interpretability and explicit reasoning required for sensitive applications such as the classification of halal meats. At the same time, the Gaussian process regression (GPR) has emerged as a powerful probabilistic framework that can model uncertainty using non-parametric Bayesian inference and provide robust predictions for large-scale multilingual text data [31]. Moreover, graphical models offer structured representations of variables and domain relationships that allow for interpretable reasoning [32], but they are still limited in capturing the underlying semantic content of non-structured input. Furthermore, ensemble and composite learning methods improve generalization performance by combining multiple models and reducing variability and bias but typically lack explicit domain integration, which limits their interpretation in regulatory settings.

Inspired by the complementary strengths and limitations of these paradigms, recent research has highlighted the integration of data-driven learning with knowledge-based thinking. In this context, the proposed XLM-R + Ontology hybrid framework combines multilingual trans-linguistic semantic coding with domain-specific classification based on the ontology, which allows both accurate and interpretable classification. Despite these advances, gaps remain, including limited public datasets, a lack of cross-linguistic regulatory mapping, and an unclear certification logic [33]. These findings suggest the need for hybrid, interpretable, and domain-aware computational frameworks integrating multilingual transformers with structured knowledge from the halal domain.

2.1. Motivation

Recent advances in machine learning have demonstrated that high-performance systems incorporate multiple uncertainty, structure, and generalization paradigms, including GPR to model uncertainty, structural models to model structure, and ensemble learning to generalize. Inspired by these principles, this paper adopts a neurosyntactic thinking approach that integrates ontological knowledge, transformational contextual learning, and fusion mechanisms to improve the resilience and decision-making in the halal food classification process.

3. Research Methodology

This study adopts a hybrid transformer–ontology framework for the advancement of the classification of halal food through multilingual ingredient labeling. The proposed methodology, as an approach, as shown in Figure 1, combines the representativity of deep learning models based on the transformer with the semantic rigor of bilingual lexical ontology to address the limitations of monolingual or purely statistical approaches to the interpretation of halal ingredients.

The proposed model has been trained using a learning rate of 16 base sizes and six epochs. In addition, the AdamW optimization was applied with parameters of type beta. The model has also been implemented in the PyTorch framework using NVIDIA hardware-based graphics, 32 GB of RAM, and an M2 processor. Experiments were performed with a hyperparameter configuration at baseline. The following section describes in detail the data collection and preprocessing, the lexical analysis using the halal ingredient ontology, the semantic representation using the transform, and the training of the classification model, as shown in Figure 1.

3.1. Research design

3.1.1. Data collection and preprocessing

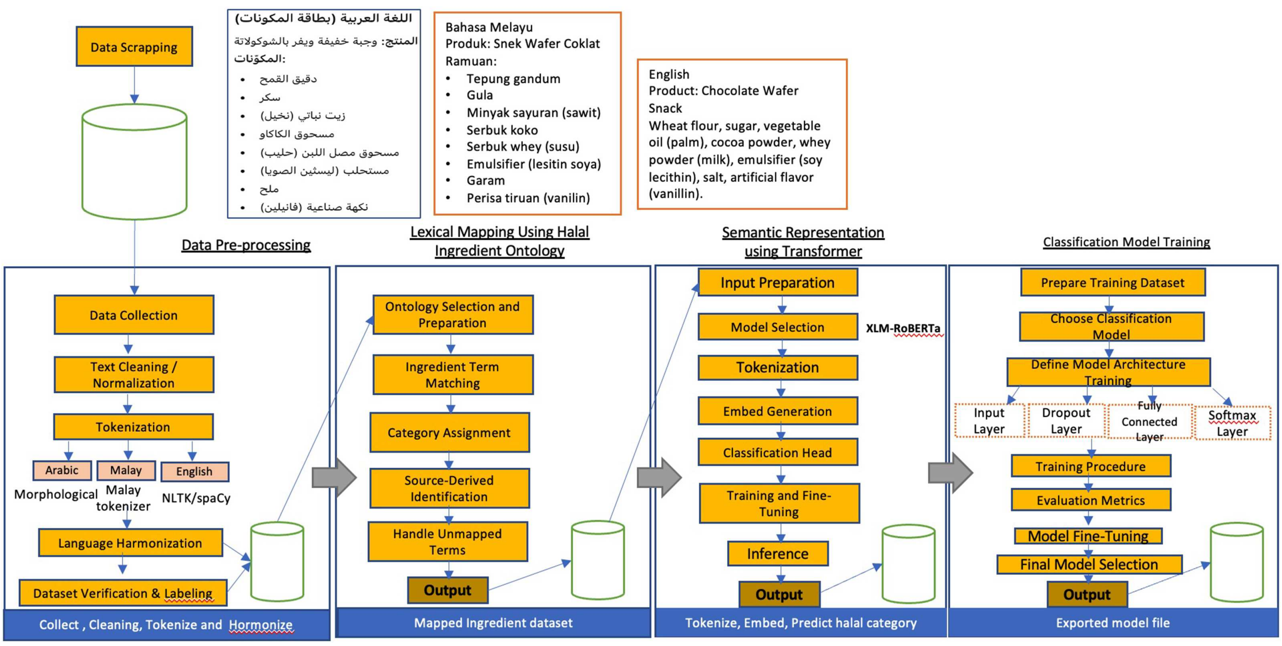
The dataset used in this study was systematically constructed by aggregating ingredient lists from multiple authoritative and heterogeneous sources, including certified halal product databases, commercial e-commerce platforms, digital food label repositories, and official manufacturer documentation. To ensure broad coverage and representativeness of the halal food domain, products were selected across diverse categories, including beverages, snacks, meat-based products, dairy items, and processed foods. Each data entry was enriched with structured metadata comprising product name, brand identifier, source, and original language context, thereby supporting traceability and reproducibility.

Raw ingredient text extracted from product labels is inherently noisy and heterogeneous, often containing hypertext markup language (HTML) artifacts, inconsistent punctuation, optical character recognition errors, and informal or abbreviated expressions. To address these issues, a rigorous preprocessing and standardization pipeline was implemented. This pipeline included character normalization, removal of non-semantic textual artifacts, unification of measurement units (e.g., mapping "g" and "grams" to a single canonical form), and orthographic correction procedures for both Malay and English texts. For Arabic ingredient descriptions, additional normalization steps were applied, including the removal of diacritical marks and orthographic unification, specifically consolidating variant alif forms (أ، إ، آ) into a standardized base character (ا), to ensure representational consistency across orthographic variants.

Following normalization, ingredient lists were subjected to structured segmentation and tokenization, decomposing compound ingredient expressions and descriptive modifiers into discrete computational units. This step enabled the precise representation of ingredient components and their attributes in a machine-readable format. Subsequently, multilingual linguistic harmonization was performed to establish semantic equivalence across Malay, English, and Arabic ingredient terminology. This process leveraged bilingual dictionaries, domain-specific glossaries, and a curated halal lexical resource to map linguistically diverse surface forms into unified canonical representations. For example, semantically equivalent expressions such as the Malay term "lemak haiwan," the English phrase "animal fat," and the Arabic term "دهن حيواني" were aligned into a single standardized representation. This harmonization mechanism is critical for reducing cross-linguistic ambiguity and enabling consistent downstream semantic analysis and classification.

The final output of this stage is a cleaned, structurally organized, and multilingual dataset that forms the foundational

Figure 1
Proposed approach



infrastructure for subsequent computational processes, including contextual embedding generation, semantic relationship modeling, and halal status classification. By enforcing linguistic homogeneity and semantic coherence at the data preparation stage, the proposed pipeline ensures that downstream machine learning and transformer-based models operate on cognitively consistent representations, thereby improving classification reliability across diverse product sources and languages.

3.1.2. Lexical mapping using halal ingredient ontology

At this stage, the canonicalized ingredient terms obtained from the preprocessing phase are systematically aligned with a formally structured halal ingredient ontology to determine both their halal classification and their biological or synthetic origin. The core classes include ingredient, sources, process, halal status, evidence, and confidence level. The ontology used in this study was derived from established halal certification knowledge bases and further extended by integrating authoritative fatwa documentation and standardized ingredient reference sources. The ontology of halal ingredients has been developed on the basis of a methodology framework that provides a structured process of specification, conceptualization, formalization, implementation, and evaluation. This ontological framework explicitly models hierarchical relationships between ingredient entities, their origin pathways, namely, plant, animal, microbial, and synthetic chemical sources, and their corresponding halal classification categories: halal (permissible), haram (prohibited), and syubhah (ambiguous or doubtful) [4].

To ensure multilingual operability and semantic consistency across heterogeneous labeling practices, the ontology was designed to incorporate terminological equivalents in Malay, Arabic, and English. This multilingual representation enables language-independent interpretation of ingredient semantics and supports consistent reasoning regardless of the source language used in product labeling [34]. Ingredient-to-ontology alignment was performed using a hybrid lexical matching strategy that combines rule-based string matching, curated synonym dictionaries, standardized naming protocols, and orthographic variation handling techniques. This multi-layered matching approach ensures robust identification of ingredients despite lexical diversity, spelling inconsistencies, transliteration variants, and cross-linguistic discrepancies.

The classification process explicitly accounts for lexical and nomenclatural variation commonly observed in food ingredient terminology. For example, terms such as “gelatin,” “hydrolyzed collagen,” and the European food additive designation “E441” were treated as semantically equivalent despite their distinct surface forms [1]. However, the determination of halal status for these functionally equivalent ingredients follows a source-dependent reasoning logic, where classification is contingent on the identified origin (e.g., bovine, porcine, aquatic, or synthetic sources) [4]. Upon successful ontological alignment, each ingredient is deterministically assigned to one of the three mutually exclusive halal categories: halal, haram, or syubhah [4].

In cases where ingredient origin information is incomplete or inherently ambiguous, such as emulsifiers or functional additives that may be derived from multiple competing sources, supplementary metadata fields are incorporated to explicitly encode potential origin distributions. Furthermore, when classification requires probabilistic reasoning rather than strict rule-based inference, a normalized confidence score is computed to quantify the degree of certainty associated with the assigned halal status.

Ingredient entries that could not be conclusively resolved through automated ontological alignment were systematically flagged for manual verification. This verification process involved cross-referencing established halal certification repositories, conducting targeted literature reviews, and, where necessary, engaging in direct communication with manufacturers. This procedure minimizes the risk of misclassification arising from lexical ambiguity, unconventional labeling practices, or transliteration inconsistencies. Subsequently, newly validated elements have been incorporated into the ontology, which has allowed for iterative improvement of the knowledge base and has supported scalability as new elements and multilingual variants emerge [34]. The output of this stage is a lexically enriched and semantically structured dataset, in which each ingredient token is linked to a standardized canonical identifier, a halal classification label, and explicit source-origin metadata. This representation provides a cognitively interpretable and computationally tractable foundation for subsequent transformer-based semantic modeling and classification.

3.1.3. Semantic representation using transformer

In this step, contextually informed semantic representation of ingredient token sequences has been systematically generated by using a multilingual encoding structure that operates on a bidirectional self-detection mechanism to capture contextual dependencies across the entire ingredient sequence, which makes it significantly easier to interpret the information in a more accurate way than traditional approaches based on a bag of words or keyword matching [35]. This architectural advantage is particularly important in ingredient-disambiguation scenarios where the same nomenclature gives substantially different halal status depending on the qualitative specification of the source, for example, “gelatin (cattle)” versus “gelatin (porcine)” [34]. The BERT multilingual encoding model has been strategically chosen for implementation based on its proven superior performance in multilingual text processing and its proven ability to encode Bengali, English, and other language formats in a unified, shared representation space [36], to enable coherent semantic processing across linguistically diverse ingredient nomenclature as found in the context of halal food labeling.

The preprocessed and lexically mapped ingredient lists were formatted as sequence text inputs, retaining the order, grouping, and source descriptors to preserve semantic integrity, and the model was fine-tuned on the basis of domain-specific training data, with each ingredient sequence being associated with a true halal classification label. The encoder was equipped with a classification head consisting of a dropout regularization and fully connected softmax layers to predict one of three permissive categories: halal, haram, or syubhah, which allows the model to learn contextual cues, including source modifiers, processing-related information, and co-occurrence patterns, indicating components of ambiguous or mixed origin. To address inherent model bias and to improve the performance of the classification, a training process following a curriculum learning strategy was carried out, where the model was trained with high-confidence-label components before a gradual inclusion of syubhah and source-dependent cases, and evaluation was done by stratified cross-validation and class-balanced performance metrics, including the macro-averaged F1 score and Matthew’s correlation coefficient. The output of step 3 is the contextual embedding generated by the transformer and the corresponding probabilistic halal classification of each ingredient sequence, which forms the basis for the integration at the system level and the final decision-making in the following steps.

3.1.4. Classification model training

In this step, a supervised learning model was developed and trained to classify the ingredient lists into halal, haram, and syubhah categories using contextual embeddings as input and pre-trained category labels as supervised learning goals, with a 70–80% training, 10–15% validation, and 10–15% testing subgroup. The fine-tuning of the transformer model with an attached classification head was preferred to conventional machine learning algorithms due to its better ability to capture contextual and multilingual relationships in the ingredient nomenclature [35]. The model architecture consists of an input layer that accepts contextual embeddings, namely, a dropout regularization layer to mitigate overfitting; a fully connected layer with three output neurons corresponding to halal, haram, and syubhah that maps embeddings to output categories; and a softmax layer that produces class distributions with validation performance-informed hyperparameter tuning [37]. Early stopping, class-balanced assessment metrics including the macro-averaged F1 score and the macro-averaged accuracy and recall, and an optional analysis of the ROC-AUC to avoid overfitting and address class differences were implemented [38]. ROC-AUC analysis helps to detect potential discrimination bias between minority classes and validates whether the model maintains a consistent discriminatory performance beyond the default threshold of Softmax. The final model, selected based on better validation performance and then exported in deployment-compatible (.pt or.h5) format, consists of a trained classifier that can accurately predict the eligibility categories for new ingredient lists in inference operations.

3.2. Fusion algorithm contribution

The full fusion score above is divided into two terms: (1) contribution weighted by the ontology and (2) contribution to transformer prediction, where the fusion algorithm combines $\hat{y}$ from the classification head (prediction based on the data) and information from the ontology $X^{\text{ontology}} = \{c_i, s_i, r_i\}$. Below is the Equation (1) score.

$$\text{Score}\,(X) = \sum w_i \rho_i + \beta.\,(\hat{y}.\,v) \tag{1}$$

Equation (1) is re-formulated as a linear fusion model inspired by the neurosymmetric integration framework, where heterogeneous sources of evidence are combined by weighted soft limits. The linear combination is chosen for its clarity and equivalence to linear models and logarithmic models commonly used in probabilistic graphical modeling. Assuming conditional independence between the ontological evidence and the transformer predictions, the combined score may be interpreted as a weighted convergence. The pseudocode of Algorithm 1 below shows the algorithm for the fusion classification of halal food.

Algorithm 1

```
FusionHalalClassification(X, Ontology, Transformer, β, θ1, θ2):
Input:
  X = list of ingredients {i1, i2, ..., in}
  Ontology = mapping of ingredient → (ci, ri)
  Transformer = trained model
  β = weighting parameter
  θ1, θ2 = decision thresholds
Output:
  y = {Halal, Haram, Syubhah}
// Step 1: Initialize ontology score
ontology_score = 0
// Step 2: Compute IDF weights
for each ingredient i in X:
wi = IDF(i) / max_IDF
         if i exists in Ontology:
            (ci, ri) = Ontology[i]
            if ci == "Halal":
              delta = +1
            else if ci == "Haram":
              delta = -1
            else:
              delta = 0
            p_i = ri * delta
         else:
            p_i = 0 # unknown ingredient
         ontology_score + = wi * p_i
//Step 3: Transformer prediction
  y_hat = Transformer.predict(X)
 transformer_score = y_hat[0] - y_hat[1]
//Step 4: Fusion
 final_score = ontology_score + β * transformer_score
//Step 5: Decision
 if final_score > θ1:
  return "Halal"
elseif final_score < θ2:
 return "Haram"
else:
 return "Syubhah"
```

Based on the pseudocode above, there are five steps to produce output:

Step 1: The ontology score is initial as 0. Here, compute ingredient contribution, which uses $\rho_i = r_i{\cdot}\delta(c_i)$, where ri = reliability of ontology classification and $\delta(c_i) = +1\delta(c_i)$, indicate +1 halal, −1 haram, and 0 syubhah. For example, computation: step 1 is to look up the ingredient in the halal ontology, such as in Table 1. In this step, each ingredient's halal status label c_i, obtained from the halal ingredient ontology and converted to a numerical value $\delta(c_i)$, as shown in Table 1. The incorporation confidence level associated with that classification, where $\boldsymbol{r_i}$ reflects the reliability of the halal status source, which originates from a combination of many sources. Thus, ρ_{i} indicates the final weighted halal-status contribution for ingredient *I*, where $\delta(c_i)$ is multiplied by the confidence $\boldsymbol{r_i}$**.** This indicates that if the ingredient is definitely haram and the confidence is high, ρ_i is strongly negative, or if the ingredient is definitely halal and the confidence is high, ρ_i is strongly positive, or if the ingredient is definitely syubhah and the confidence is low, ρ_i is strongly negative.

Step 2: To compute the ingredient's important weight, $w_{i,}$ based on the inverse document frequency (IDF) from information retrieval theory. The weights w are computed using an IDF scheme to down-weight common ingredients and emphasize rare but discriminative ingredients.

Each ingredient obtains a Semantic Importance Weight as in Equation (2):

$$w = \frac{IDF_{(i)}}{\max(IDF)} \tag{2}$$

IDF (i) = log (N/df_i), where N is the total number of products in the dataset and df_i is the number of products containing

Table 1
Computation steps, conversion, and confidence level associated with the classification of the ingredient halal ontology

Ingredient	Source	Halal status c_i	Confidence r_i	$\delta(c\_i)$	r_i	$\rho_i = r_i.\ \delta(c\_i)$
Gelatin	Porcine	Haram	1.0	–1	1.0	–1.0
Lecithin	Soy	Halal	0.95	+1	0.95	+0.95
Flavoring	Unknown	Syubhah	0.5	0	0.5	0.0

the ingredient i. If an ingredient appears in many products, dfi is high IDF is low ingredient, and it is not discriminative. However, if an ingredient appears in a few products, it indicates that the IDF is high. Therefore, the ingredient is important for classification. Rare ingredients that appear in a few products get a higher weight, indicating more important for determining halal/haram. Table 2 shows the calculation of the important weight. Common ingredients (like "sugar" or "water") get a lower weight indicating less impact on classification. For example, N = 1000 products in the dataset, Ingredient i = gelatin appears in 25 products, IDF (gelatin) = log 1000/ 25 = 3.69, log (1000/25) = 3.69, Ingredient j = lecithin appears in 120 products, and IDF (lecithin) = log 1000/ 120 = 2.12. The normalization for lecithin is 2.12/3.69 = 0.68. In these first terms, weight w_i is multiplied by ρ_i as the first ontology-derived score that captures rule-based knowledge, but the ingredients list is often ambiguous, abbreviated, or context-dependent.

Step 3: The semantic and contextual meanings are integrated through the transformer classifier's output. Therefore, the second term, called the transformer prediction, is constructed using the following formula: β. ($\hat{y}$. v). The parameter β is considered to be a tunable hyperparameter that controls the ratio between symbolic (ontological) and neural (transcriptive) evidence. β is optimized via grid search over [0,1] with step size 0.1 using validation accuracy.

The transformer produces the probability vector, $\hat{y}$ = [pHalal, pHaram, pSyubhah]. In contrast, each element is obtained from the softmax classification head, $\hat{y}$ = Softmax (Wh + b), where h is the final hidden representation of the ingredient sequence. At the same time, W and b are learnable parameters of the classification layer. To integrate the model's prediction into the score framework, $\hat{y}$ is mapped to a single continuous value via a class-weight vector, v = [1, −1, 0]. Thus, the transformer_score = y_hat[0] - y_hat[1] is executed. The conflicts between ontology evidence and transformer predictions are resolved through weighted competition. When ontology evidence is strong (high $\Sigma w_i \rho_i$), it dominates the decision. Otherwise, the transformer prediction prevails. This soft resolution avoids hard rule overrides and improves robustness in ambiguous cases. In contrast, element 1 indicates in class halal (positive, stronger halal evidence), element −1 indicates in class haram (negative, stronger haram evidence), and 0 indicates class syubhah (uncertainty, no direction bias), $\hat{y}$.v = pHalal−pHaram. To balance this, the scaling parameter β controls the influence of the transformer model relative to ontology reasoning, β($\hat{y}$·v). The balancing β is high, indicating the system relies on data-driven learning, while β is low, indicating the system prioritizes halal certification knowledge. For example, interpretation of transformer classifier output, suppose the $\hat{y}$ = [0.65,0.25,0.10], $\hat{y}$·v = 0.65−0.25 = 0.40. Given β is 0.8, thus β($\hat{y}$·v), 0.8 (0.4) = 0.32. The final product score is positive evidence.

Table 2
Calculation of the ingredient's important weight, w_i, based on inverse document frequency (IDF)

Ingredient	df_i	IDF(i) = log (N/df_i)	Normalized w_iIDF(i)/ max (IDF)
Flavoring	500	Log (1000/500) ≈ 0.02	0.01
Lecithin	120	Log (1000/120) ≈ 2.12	0.68
Gelatin	25	Log (1000/25) ≈ 3.69	1.00

Step 4: The combination of OntologyScore(X) and $\hat{y}$ (transformer prediction) to the score, Score (X) = Σ w_i ρ_i + β. ($\hat{y}$. v). The threshold-based decision function is applied to assign the final halal class label, $r1$ > 0, typically small for a + 0.2, and $r2$ < 0, typically symmetric small for a −0.2. The score is interpreted along a continuous spectrum, where positive values indicate a halal tendency, negative values indicate a haram tendency, and values near zero represent uncertainty. For example, consider the following calculation. The first ontology score is calculated by using OntologyScore(X) = Σ w_i ρ_i = (0.01)(0.0) + (0.68)(0.95) + (1.00)(−1.0) = −0.35. Next, the transformer semantic prediction output by using β($\hat{y}$·v) = 0.32, indicating moderate semantic evidence supporting the halal class. Given the decision thresholds $r1$ = +0.20 and $r2$ = −0.20, the resulting score falls within the uncertainty interval ($r2$ ≤ Score(X) ≤ $r1$). Thus, Score (X) = −0.35 + 0.32 = −0.034. Since the result score (x) < $r2$, the product is classified as *syubhah*.

4. Result and Discussion

The experimental evaluation used a curated dataset of 10,000 records of food products collected from multiple sources, including e-commerce platforms, halal-certified and ingredient databases, and food labeling repositories, where it is classified into three halal status classes: syubhah (3400 samples), haram (3300 samples), and halal (3300 samples), as shown in Table 3. The dataset is semi-parallel in nature, which are 65% of records containing aligned translations (Malay–English or Malay–Arabic). In comparison, another 35% are independently sourced monolingual entries, later aligned at the ingredient level using lexical mapping.

Before training the model, the dataset has undergone rigorous preprocessing to ensure data integrity and avoid information leakage. Duplicates have been eliminated, and ingredient descriptions have been carefully screened to exclude any explicit class-indicative terms, thus reducing the risk that models learn to respond to trivial label signals instead of meaningful semantic patterns.

The dataset was then split into training and testing subsets by a stratified 80:20 split, maintaining a proportional distribution of each class in both sets [39]. The resulting training set consisted of 2720 syubhah, 2640 haram, and 2640 halal samples, while the test set consisted of 680 syubhah, 660 haram, and 660 halal samples. Additional overlap checks verified that no products were found in either the training or the test sets, which ensured that

Table 3
Dataset distribution

Attribute	Category	Count	Percentage (%)
Language	Malay	10,000	100
	English	10,000	100
	Arabic	10,000	100
Record type	Parallel	6500	65
	Non-parallel	3500	35
	Non-parallel	3500	35
Class label	Halal	3400	34
	Haram	3300	33
	Syubhah	3300	33
Geographic source	Malaysia	60%	—
	Middle East	25%	—
	Global	15%	—

the performance assessment accurately reflected the models' ability to generalize to unobservable data. This careful preparation of the dataset provides a reliable basis for a fair comparison between traditional machine learning classifiers, deep learning architectures, and transformer-based models for the task of classifying halal ingredients.

The training and validation performance of the model proposed over six epochs is summarized in Table 4. A rapid reduction in the loss of training between epoch 1 and epoch 2 was observed, from 0.2500 to 0.1571. This substantial improvement shows the effectiveness of the proposed learning framework in capturing discriminatory characteristics early in the training process.

From the second epoch onward, the loss of training gradually declined until it reached 0.1519 in epoch 6. However, the accuracy of both the validation and the F1 score remained stable at 87.35% and 0.8769, respectively. This stability indicates that the model reached convergence early and maintained a consistent generalization performance for the non-visible validation data. Importantly, despite continued training, no degradation of validation performance was observed, which suggests that no overfitting occurred. Overall, the results confirm that the proposed model is performing effectively and that it demonstrated rapid learning capability with a significant reduction in training loss, stable accuracy, and F1 score, resulting in effective extraction of discriminative semantic features from multilingual halal ingredient data.

The identical accuracy and F1 score values observed across epochs 2–6 are attributed to early convergence of the model, where performance stabilized after reaching a local optimum.

Table 4
Training and validation performance

Epoch	Training loss	Validation accuracy	Validation F1 score
1	0.2500	87.35	0.8769
2	0.1571	87.35	0.8769
3	0.1552	87.35	0.8769
4	0.1538	87.35	0.8769
5	0.1526	87.35	0.8769
6	0.1519	87.35	0.8769

Early stopping criteria were applied, and minimal parameter updates occurred beyond epoch 2 due to low gradient variance and stabilized loss function behavior. This indicates model convergence rather than evaluation error.

4.1. Model performance comparison

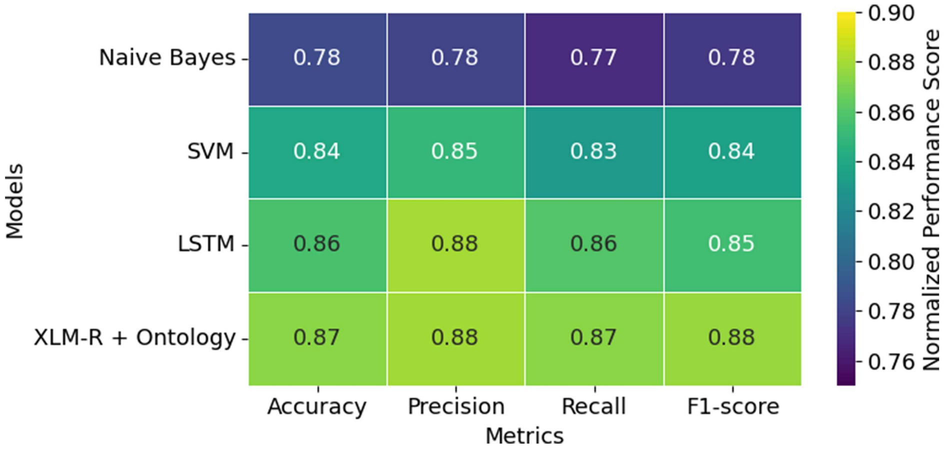
The evaluation of halal food classification models shows a clear difference between the conventional base models and the proposed hybrid model. Figure 2 summarizes the performance metrics for Naive Bayesian (NB), Support Vector Machine(SVM), Long Short-Term Memory (LSTM), and Cross Lingual Model with Roberta (XLM-R) + Ontology model in terms of accuracy, recall, and F1 score for the halal, haram, and syubhah datasets. Based on Figure 2, the comparative performance of traditional machine learning models, deep learning architectures, and transform-based approaches for the halal–haram–syubhah classification challenge reveals clear differences in efficiency and robustness. The NB classifier showed the lowest overall performance of the evaluated models, with an accuracy of 0.784 and an F1 score of 0.775. Although accuracy remained low (0.781), the lower recall (0.769), as shown in Figure 2, indicates a limited ability to consistently identify class-specific patterns within the three halal categories. This behavior reflects the well-known assumption of independence of the NB, which limits its performance on complex semantic classification tasks involving ingredient lists and contextual cues [40, 41].

SVMs showed an improvement over NB with an accuracy of 0.840 and an F1 score of 0.836. The SVM achieved a more balanced balance between accuracy (0.842) and recall (0.831), which suggests an improvement in the level of discrimination between the halal, haram, and syubhah classes. However, its performance was still limited by its reliance on fixed-item representations, which limited its ability to capture the deeper semantic relationships that are present in ingredient descriptions.

The LSTM model further improved the classification performance, reaching an accuracy of 0.857 and an F1 score of 0.854. The model made use of its sequential modeling capability, which allowed for a better contextual understanding than traditional classifiers. By processing ingredient lists as ordered sequences, the LSTM learns the contextual dependencies between the ingredients, reducing ambiguities arising from the names of the ingredients and improving the consistency of the classification between the *halal*, *haram*, and *syubhah* categories. However, its performance has plateaued compared to transformer methods, probably due to a lack of extensive training and limited ability to capture long-term semantic dependencies in multilingual and transliterated ingredient data.

The hybrid XLM-R + Ontology model consistently outperformed all the baseline approaches, achieving the highest accuracy (0.874) and F1 score (0.877), as well as higher precision (0.879) and recall (0.874). Strong and balanced performance across all assessment metrics underlines the benefits of integrating transformer-based contextual embeddings with structured ontology. This hybrid neural–symbolic framework allows for a more precise semantic interpretation of ingredient terms, a robust handling of ambiguous syubhah cases, and a reliable identification of haram indicators, which improves the performance of generalization [42]. Overall, the results in Figure 2 confirm that while traditional and sequential deep learning models provide incremental improvements, the proposed hybrid XLM-R + ontology approach is the most efficient and stable solution for the classification of halal ingredients, especially in complex, domain-specific, and multilingual environments.

Figure 2
Performance metric heatmap of the classification model

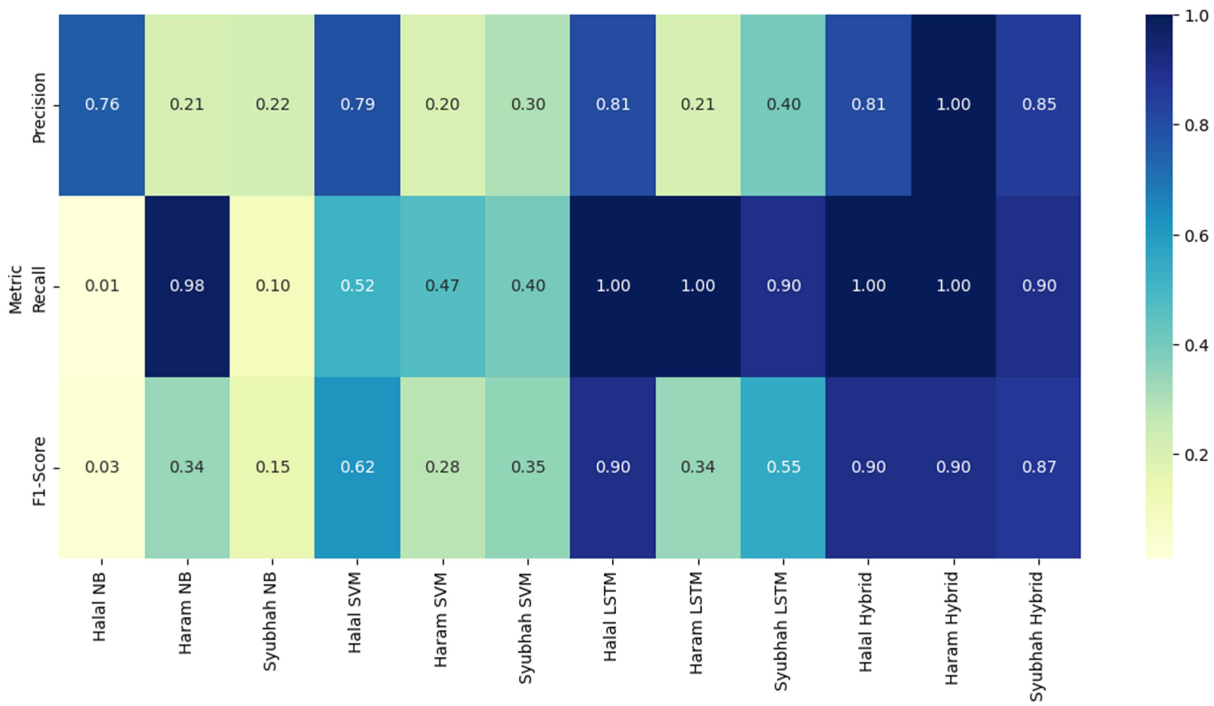


4.1.1. Attention map

Figure 3 shows a comparative assessment of the four classifiers, namely, NB, SVM, LSTM, and XLM-R + Ontology, highlighting significant differences in their ability to classify halal, haram, and syubhah products. NB shows the weakest performance, with a general failure to compare class predictions. Although it achieves high accuracy for halal (0.76), the exclusion is almost negligible (0.01), indicating extreme under-reporting of halal cases. Conversely, haram shows a high recall (0.98) but a low accuracy (0.21), indicating frequent false-positive events. Detection of syubhah residues shows a slight recall (0.10) but a low accuracy (0.22), reflecting poor generalization across all classes. The macro-averaged F1 score for NB is only 0.16, confirming that the Bayes system is inefficient in general.

The SVM provides a slight improvement and achieves a more balanced performance between grades. The accuracy for halal, haram, and syubhah is 0.79, 0.20, and 0.30, while the recall values are 0.52, 0.47, and 0.40. This results in a macro-averaged F1 score of 0.42, indicating that SVM partially captures semantic differences but still incorrectly classifies haram and syubhah instances due to the linearity constraints of the TF-IDF Feature space. The LSTM shows a high accuracy for halal and haram (1.0 each) and syubhah (0.90) but a low accuracy for haram (0.21) and syubhah (0.40), indicating a strong tendency to overestimate. The macro-averaged F1 score of 0.60 indicates that although the LSTM identifies the most positive cases, it is struggling with specificity, probably due to the limited data and lack of integrated knowledge of the field.

Figure 3
Attention map of performance metric heatmap of classification model (Halal-Haram dataset)



The XLM-R + Ontology hybrid model significantly outperforms the baselines, reaching a precision of 0.635 and a macro-F1 score of 0.89. Halal and haram detection achieve a perfect recall (1.0) with a high accuracy (0.81 and 1.00, respectively), giving 0.90 F1 scores for both classes. Syubhah is detected with a precision of 0.85, a recall of 0.90, and an F1 score of 0.87, indicating robust performance in all categories. This shows that the combination of transformer-based embeddings with ontology rules effectively captures the semantics and context of the ingredients and solves many of the ambiguities that hamper traditional and LSTM models.

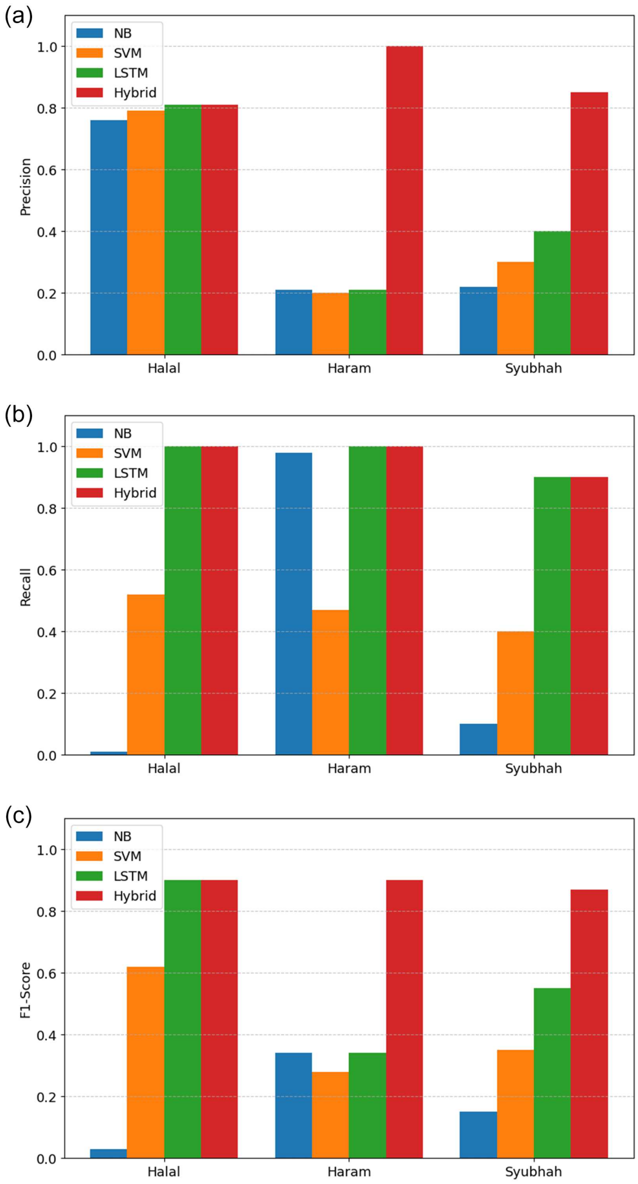
4.1.2. Performance distribution map

A comparative analysis of NB, SVM, LSTM, and XLM-R + Ontology (hybrid) models reveals significant differences in their ability to classify halal, haram, and syubhah products, as shown in Figure 4. As shown in the combined performance distribution graph, the hybrid model has consistently outperformed the traditional classifiers in all metrics. In particular, the NB shows a moderate accuracy for halal (0.76) but an extremely low accuracy for positive cases (0.01), indicating a serious under-detection of positive cases, especially in minority and ambiguous classes such as syubhah (0.10, F1 score 0.15).

As shown in Figure 4, SVM shows a slight improvement, with improved recall and F1 scores for halal (0.52 and 0.62), but performance remains limited for haram and syubhah, which highlights the limitations of balancing accuracy and recall in an unbalanced dataset. LSTM achieves a high recall for halal and haram (1.00), but inconsistent accuracy (0.21 for haram) results in an uneven F1 score, which may reflect an overestimation bias, possibly due to insufficient data coverage or a limited multilingual coverage. The XLM-R + Ontology hybrid model. On the other hand, it achieves the highest and most balanced performance in all grades, with accuracy ranging from 0.81 to 1.00, recall from 0.90 to 1.00, and F1 scores from 0.85 to 0.90, demonstrating its ability to integrate contextual multilingual learning with structured domain knowledge.

This hybrid framework effectively captures semantic nuances and domain-specific indicators, which significantly improve the detection of low-resource classes such as haram and syubhah. Overall, the findings highlight the superiority of neural–symbolic hybrid approaches for complex, multilingual, and domain-specific classification tasks and confirm that integration of contextual embedding with knowledge-based rules significantly increases the reliability and robustness of the classification at the product level.

Figure 4
(a) Precision distribution, (b) recall distribution, and (c) F1 score distribution

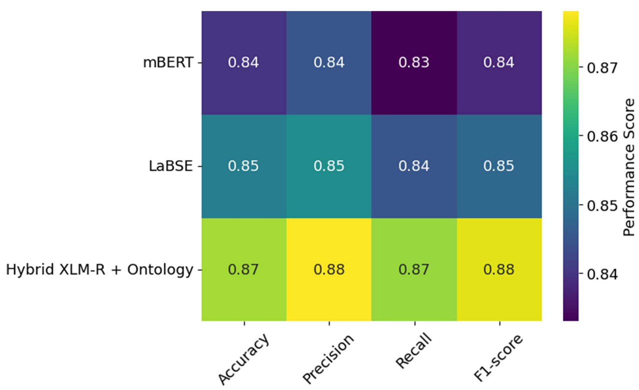


4.2. Model performance comparison of another multilingual transformer

For further evaluation, this proposed model was also compared to another multilingual transformer.

Based on Figure 5, the proposed model, hybrid XLM-R with ontology, performs well across evaluation metrics. Hybrid XLM-R with ontology model demonstrated the highest performance F1 score, 0.877, compared with mBERT and LaBSE. Other than that, in terms of accuracy, precision, and recall, the proposed approach outperforms other multilingual transformer models.

Figure 5
Performance comparison with multilingual transformers



4.3. Discussion

The results highlight the limitations of traditional classifiers in dealing with data from multiple languages and contextually ambiguous ingredients. NB showed very low recall across all minority groups, especially halal (0.01) and syubhah (0.10), although halal was slightly more accurate (0.76). This shows that simple probabilistic models struggle to capture the semantic complexity inherent in transliterated or compound terms. SVM achieved some improvements, balancing accuracy and recall more efficiently across halal, haram, and syubhah, but still failed to reliably identify minority groups, which reflects the lack of linear separation for complex semantic patterns.

Sequential models such as LSTM have also shown significant weaknesses. Although the recall values were high for halal (1.0) and haram (1.0), the accuracy was low (0.81 for halal and 0.21 for haram), indicating an overestimation of the prediction. Syubhah's predictions were relatively accurate (precision 0.40, recall 0.90), but overall instability suggests that sequencing alone is not sufficient when data are sparse and not fully verbalized, as in the case of multilingual ingredient lists and transliterations.

The XLM-R + Ontology model significantly outperformed all baselines and achieved high accuracy and recall in all classes: halal (0.81, recall 1.0, F1 0.90), haram (0.85, recall 1.0, F1 0.90), and syubhah (0.87). This shows the effectiveness of integrating symbolic knowledge into transformer-based embeddings. XLM-R effectively captured deep contextual and multilingual semantics, and the ontology layer encoded explicit haram-like descriptors like gelatin and ethanol. The synergies between these components have allowed for robust performance in both majority and minority classes and have resolved ambiguities that hamper traditional classifications.

Analysis of the attention map, as shown in Figure 3, provides further insight into the behavior of the model. For halal products, the model focuses on semantically significant ingredients such as water, milk, and sugar, which are correlated with high recall and accurate detection. In the case of haram, the emphasis is on symbols such as gelatin and ethanol, which reflect the successful integration of the ontological rules. The distribution of attention in syubhah is more diffuse, indicating inherent semantic ambiguity and partly explaining the lower recall rate. Overlapping performance metrics on attention maps further reveal the source of accuracy errors (false positives) and memory failures (false negatives) and reveal areas for improvement, such as extending the syubhah-labeled examples or the ontology coverage.

These trends are confirmed by the performance maps as shown in Figure 4. NB consistently underestimates non-halal and syubhah compounds, SVM shows partial improvement with moderate accuracy, and LSTM has low precision and unstable recall. On the other hand, the XLM-R + Ontology hybrid framework achieves better and more balanced performance across metrics, which shows the benefits of combining deep contextualization with structured domain knowledge. Despite these improvements, class imbalance remains a problem, particularly for syubhah detection, suggesting that further strategies such as data augmentation, synthetic sampling, or loss-aware class function could further improve reliability. Taken together, these findings highlight the need for a hybrid neural–symmetric approach to the accurate, comprehensible, and multilingual classification of halal foods.

The proposed model results, as in Figure 5, show stronger performance in handling classes of halal, haram, and syubhah compared to other hybrid models. This combination transformer-based with domain reasoning provides important reasoning for domain-specific classification tasks, for example, in halal compliance. However, the deployment of automated halal classification systems introduces critical ethical considerations, particularly the risk of incorrect classification leading to unintended religious non-compliance. A false "halal" prediction in sensitive cases may have significant social and religious implications.

5. Conclusion

This study has introduced a hybrid neural–symbolic halal ingredient classification framework explicitly integrating a multilingual transformation model (XLM-R) with domain-based reasoning to address the problems of semantic ambiguity, multilingual ingredient descriptions, and interpretability requirements in the assessment of halal compliance. Under the proposed architecture, the transformer acts as a basic semantic encoder, enabling robust learning of cross-lingual representations, while the ontology provides clear, rule-based, halal knowledge that supports transparent and trust-based decision-making. The experimental results show that the hybrid XLM-R + Ontology model consistently outperforms traditional machine learning classifiers and deep learning models from LSTM across all assessment metrics. The ability of the transformer to capture contextual and cross-linguistic semantics, combined with ontological-guided reasoning, gives a particularly strong performance in identifying syubhah and haram components accurately, two of the most critical and error-prone categories in halal assessment. Attentional and performance attribution analyses further confirm that synergies between transformation and ontology enhance class-wise discrimination and reduce semantic misclassification.

From an applied perspective, the proposed framework offers a scalable, interpretable, and linguistically robust solution for deployment in halal certification authorities, food industry compliance systems, e-commerce platforms, and consumer-facing applications. Beyond halal food classification, the presented hybrid transformer-based approach is extensible to other regulatory and compliance-driven domains that require multilingual text understanding combined with structured expert knowledge.

Future work will explore extending the halal ontology with region-specific fatwas, integrating explainable AI modules to provide fine-grained justifications of transformer predictions, and incorporating continual learning mechanisms to adapt the transformer to evolving ingredient formulations and emerging food additives. These directions aim to further strengthen the reliability, transparency, and long-term applicability of a hybrid transformer-based neural–symbolic system.

Acknowledgment

We would like to extend our thanks to Universiti Teknikal Malaysia Melaka (UTeM) and Fakulti Teknologi Maklumat dan Komunikasi (FTMK) for their support.

Funding Support

This work is sponsored by Universiti Teknikal Malaysia Melaka (UTeM), Grant Number: URMG-AJMAN/2024/FTMK/A00070.

Ethical Statement

This study does not contain any studies with human or animal subjects performed by any of the authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

The data that support the findings of this study are openly available at https://doi.org/10.5281/zenodo.19173858.

Author Contribution Statement

Mohd Azmi Al Betar: Conceptualization, Methodology, Supervision. **Noorrezam Yusop:** Conceptualization, Methodology, Software, Formal analysis, Resources, Writing – original draft, Writing – review & editing, Project administration. **Tao Hai:** Methodology, Validation, Software, Supervision, Project administration, Investigation. **Nor Aiza Moketar:** Validation, Formal analysis, Writing – review & editing. **Nuridawati Mustafa:** Investigation. **Mohd Nazrien Zaraini:** Software, Data curation, Resources. **Massila Kamalrudin:** Supervision.

References

[1] Sunmola, F., Baryannis, G., Tan, A., Co, K., & Papadakis, E. (2025). Holistic framework for blockchain-based halal compliance in supply chains enabled by artificial intelligence. *Systems*, *13*(1), 21. https://doi.org/10.3390/systems13010021

[2] Harati, A., & Farzaneh, P. (2024). The advancement in global halal industry from farm to fork. *Journal of Halal Product and Research*, *7*(2), 113–124. https://doi.org/10.20473/jhpr.vol.7-issue.2.113-124

[3] Akbar, J., Gul, M., Jahangir, M., Adnan, M., Saud, S., Hassan, S., Nawaz, T., & Fahad, S. (2023). Global trends in halal food standards: A review. *Foods*, *12*(23), 4200. https://doi.org/10.3390/foods12234200

[4] Tarannum, S., Jalal, M. S., & Huda, M. N. (2024). HALALCheck: A multi-faceted approach for intelligent halal packaged food recognition and analysis. *IEEE Access*, *12*, 28462–28474. https://doi.org/10.1109/ACCESS.2024.3367983

[5] Pansare, P. P., Khatik, K. N., Shigvan, N. N., & Lande, R. V. (2024). A review on recipe generation from food image using machine learning. *International Journal of Advanced Research in Science, Communication and Technology*, *4*(3), 271–275. https://doi.org/10.48175/IJARSCT-22241

[6] S, P., I, S., L, K., Bhoomika, P., Tejaswi, K., Khande, L. S., & Vemishett, N. S. (2024). Automating nutritional claim verification: The role of OCR and machine learning in enhancing food label transparency. In *2024 International Conference on IoT Based Control Networks and Intelligent Systems*, 1164–1171. https://doi.org/10.1109/ICICNIS64247.2024.10823177

[7] Varma, M. N., Yadav, C. S. B., Onapakala, K., Mathamsetty, A., Saibaba, C. M. H., Kumari, P. L., . . . , & Abhilash, P. (2024). Advanced transformer-based deep learning techniques for enhancing contextual understanding in natural language processing. *African Journal of Biomedical Research*, *27*(4s), 5153–5161. https://doi.org/10.53555/AJBR.v27i4S.4562

[8] Yusop, N., Kamalrudin, M., Mustafa, N., Hai, T., Zaraini, M. N., Hakimi, H., & Sardikan, S. F. N. (2025). Recent integrating machine learning and Malay-Arabic lexical mapping for halal food classification. *International Journal of Advanced Computer Science and Applications*, *16*(10). https://doi.org/10.14569/IJACSA.2025.0161067

[9] Conneau, A., Khandelwal, K., Goyal, N., Chaudhary, V., Wenzek, G., Guzmán, F., . . . , & Stoyanov, V. (2020). Unsupervised cross-lingual representation learning at scale. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 8440–8451. https://doi.org/10.18653/v1/2020.acl-main.747

[10] Conneau, A., Wu, S., Li, H., Zettlemoyer, L., & Stoyanov, V. (2020). Emerging cross-lingual structure in pretrained language models. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 6022–6034. https://doi.org/10.18653/v1/2020.acl-main.536

[11] Chaturvedi, A., Sharma, C., & Chaturvedi, M. (2025). Online shopping website review system using predictive AI. *International Journal of Scientific Research in Engineering and Management*, *9*(5), 1–9. https://doi.org/10.55041/IJSREM48588

[12] Nawaz, A., Afzal, A., Khatibi, A., Shankar, A., Madan, H., Faisal, H. S., . . . , & Islam, F. (2025). Role of artificial intelligence in halal authentication and traceability: A concurrent review. *Food Control*, *169*(111003). https://doi.org/10.1016/j.foodcont.2024.111003

[13] Kulsum, A. S., Abhisha, B. H., K, B., & A, S. (2025). NutriSafe: Allergen detection using NLP and Gen AI. In *2025 International Conference on Biomedical Engineering and Sustainable Healthcare*, 1–6. https://doi.org/10.1109/ICBMESH66209.2025.11182261

[14] Tarannum, S. (2023). *Halal food identification from product ingredients using machine learning*. Master's Thesis, United International University.

[15] Mazli, M. S. (2020). *Deep learning-based halal food recognition*. Master's Thesis, Universiti Teknologi Malaysia.

[16] Lachraf, R., Nagoudi, E. M. B., Ayachi, Y., Abdelali, A., & Schwab, D. (2019). ArbEngVec : Arabic-English cross-lingual word embedding model. In *Proceedings of the Fourth Arabic Natural Language Processing Workshop*, 40–48. https://doi.org/10.18653/v1/W19-4605

[17] Mohamad Hashim, S. F., Salim, J., Mohd Noah, S. A., & Wan Mustapha, W. A. (2023). Ontology-based traceability system for halal status of flavour: A conceptual framework. *Malaysian Journal of Information and Communication Technology*, *8*(2). https://doi.org/10.53840/myjict8-2-97

[18] Courtois, M., Ostendorff, M., Hennig, L., & Rehm, G. (2024). Symmetric dot-product attention for efficient training of BERT language models. In *Findings of the Association for Computational Linguistics: ACL 2024* (pp. 8002–8011). https://doi.org/10.18653/v1/2024.findings-acl.476

[19] Kale, T. V., & Mendhe, S. (2025). A review on advances in sentiment analysis: A deep learning approach using transformer based models. In *2025 4th International Conference on Sentiment Analysis and Deep Learning*, 235–239. https://doi.org/10.1109/ICSADL65848.2025.10933230

[20] Akhila, N., P, S. J., & P, S. K. (2023). Comparative study of Bert models and Roberta in transformer based question answering. In *2023 3rd International Conference on Intelligent Technologies*, 1–5. https://doi.org/10.1109/CONIT59222.2023.10205622

[21] Alshanqiti, A. M., Albouq, S., Alkhodre, A. B., Namoun, A., & Nabil, E. (2022). Employing a multilingual transformer model for segmenting unpunctuated Arabic text. *Applied Sciences*, *12*(20), 10559. https://doi.org/10.3390/app122010559

[22] Al-Duwais, M., Al-Khalifa, H., & Al-Salman, A. (2024). A benchmark evaluation of multilingual large language models for Arabic cross-lingual named-entity recognition. *Electronics*, *13*(17), 3574. https://doi.org/10.3390/electronics13173574

[23] Patwardhan, N., Marrone, S., & Sansone, C. (2023). Transformers in the real world: A survey on NLP applications. *Information*, *14*(4), 242. https://doi.org/10.3390/info14040242
[24] Minaee, S., Kalchbrenner, N., Cambria, E., Nikzad, N., Chenaghlu, M., & Gao, J. (2022). Deep learning–based text classification: A comprehensive review. *ACM Computing Surveys*, *54*(3), 62. https://doi.org/10.1145/3439726
[25] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, *1*, 4171–4186. https://doi.org/10.18653/v1/N19-1423
[26] Qiu, X., Sun, T., Xu, Y., Shao, Y., Dai, N., & Huang, X. (2020). Pre-trained models for natural language processing: A survey. *Science China Technological Sciences*, *63*(10), 1872–1897. https://doi.org/10.1007/s11431-020-1647-3
[27] Shen, Y., Bao, W., Gao, G., Zhou, M., & Zhao, X. (2024). Unsupervised multilingual machine translation with pretrained cross-lingual encoders. *Knowledge-Based Systems*, *284*, 111304. https://doi.org/10.1016/j.knosys.2023.111304
[28] Hashmi, E., Yayilgan, S. Y., & Shaikh, S. (2024). Augmenting sentiment prediction capabilities for code-mixed tweets with multilingual transformers. *Social Network Analysis and Mining*, *14*(1), 86. https://doi.org/10.1007/s13278-024-01245-6
[29] Dong, J. (2024). Natural language processing pretraining language model for computer intelligent recognition technology. *ACM Transactions on Asian and Low-Resource Language Information Processing*, *23*(8), 110. https://doi.org/10.1145/3605210
[30] Jin, B., & Xu, X. (2025). China commodity price index (CCPI) forecasting via the neural network. *International Journal of Financial Engineering*, *12*(3), 2550003. https://doi.org/10.1142/S2424786325500033
[31] Jin, B., & Xu, X. (2025). Forecasts of coking coal futures price indices through Gaussian process regressions. *Mineral Economics*, *38*(1), 203–217. https://doi.org/10.1007/s13563-024-00472-9
[32] Jin, B., & Xu, X. (2025). A study of contemporaneous residential real estate price causation across major Jiangsu Province cities: Methodology using vector error-correction models and directed acyclic graphs. *Economics Open*, *1*, 2550008. https://doi.org/10.1142/S308284142550008X
[33] Lankford, S., Afli, H., & Way, A. (2023). adaptMLLM: Fine-tuning multilingual language models on low-resource languages with integrated LLM playgrounds. *Information*, *14*(12), 638. https://doi.org/10.3390/info14120638
[34] Liu, Y., Min, W., Jiang, S., & Rui, Y. (2024). Convolution-enhanced Bi-Branch Adaptive Transformer with cross-task interaction for food category and ingredient recognition. *IEEE Transactions on Image Processing*, *33*, 2572–2586. https://doi.org/10.1109/TIP.2024.3374211
[35] Lo, C.-K., & Simard, M. (2019). Fully unsupervised cross lingual semantic textual similarity metric based on BERT for identifying parallel data. In *Proceedings of the 23rd Conference on Computational Natural Language Learning*, 206–215. https://doi.org/10.18653/v1/K19-1020
[36] Wadud, M. A. H., Mridha, M. F., Shin, J., Nur, K., & Saha, A. K. (2023). Deep-BERT: Transfer learning for classifying multilingual offensive texts on social media. *Computer Systems Science and Engineering*, *44*(2), 1775–1791. https://doi.org/10.32604/csse.2023.027841
[37] Nair, A. R., Singh, R. P., Gupta, D., & Kumar, P. (2024). Evaluating the impact of text data augmentation on text classification tasks using DistilBERT. *Procedia Computer Science*, *235*, 102–111. https://doi.org/10.1016/j.procs.2024.04.013
[38] Pavlyshenko, B., & Stasiuk, M. (2024). Data augmentation in text classification with multiple categories. *Electronics and Information Technologies*, (*25*), 67–80. https://doi.org/10.30970/eli.25.6
[39] Sadaiyandi, J., Arumugam, P., Sangaiah, A. K., & Zhang, C. (2023). Stratified sampling-based deep learning approach to increase prediction accuracy of unbalanced dataset. *Electronics*, *12*(21), 4423. https://doi.org/10.3390/electronics12214423
[40] Araf, I., Idri, A., & Chairi, I. (2024). Cost-sensitive learning for imbalanced medical data: A review. *Artificial Intelligence Review*, *57*(4), 80. https://doi.org/10.1007/s10462-023-10652-8
[41] Kowsari, K. S., Meimandi, K. J., Heidarysafa, M., Mendu, S., Barnes, L., & Brown, D. (2019). Text classification algorithms: A survey. *Information*, *10*(4), 150. https://doi.org/10.3390/info10040150
[42] Bhuyan, B. P., Ramdane-Cherif, A., Tomar, R., & Singh, T. P. (2024). Neuro-symbolic artificial intelligence: A survey. *Neural Computing and Applications*, *36*(21), 11757–11792. https://doi.org/10.1007/s00521-024-09960-z

How to Cite: Al Betar, M. A., Yusop, N., Hai, T., Moketar, N. A., Mustafa, N., Zaraini, M. N., & Kamalrudin, M. (2026). A Hybrid Transformer–Ontology Framework for Halal Food Classification Using Multilingual Ingredient Label Analysis. *Journal of Computational and Cognitive Engineering*. https://doi.org/10.47852/bonviewJCCE62029040