J. Integr. Sci. Technol. 2025, 13(4), 1079

Journal of Integrated
SCIENCE & TECHNOLOGY

Advancing Vertigo diagnosis with large language models: A multimodal, AI-driven approach to Etiology differentiation

Mohammad Alshraideh[1,7*], Yara Alkayed[2], Abedalrahman Alshraideh[3], Yasmin Al Trabsheh[4], Bahaaldeen Alshraideh[5], Heba Alshraideh[6], Bayan Alfayoumi[7], Njwan Alshraideh[6]

[1]Artificial Intelligence Department, The University of Jordan, Amman 11942 Jordan. [2]Internal Medicine Department, The University of Jordan, Amman 11942 Jordan. [3]Internal Medicine, East Midlands Deanery, NHS, England, UK, [4]Clinical Attache, United Lincolnshire Hospitals, NHS, England, UK. [5]Division of Urology, Department of Special Surgery, The University of Jordan, Amman 11942. Jordan. [6]Medicine School, The University of Jordan, Amman, Jordan, [7]Information Technology, Lusail University, Lusail, Qatar.

Submitted on: 11-Nov-2024, Accepted and Published on: 24-Jan-2025 Article

ABSTRACT

Vertigo is a frequent and often debilitating condition arising from various underlying causes, such as vestibular disorders, neurological conditions, and systemic diseases. Accurate differentiation of its etiology is challenging due to overlapping clinical presentations. This research employs an AI-driven approach, utilizing Large Language Models (LLMs) like GEMMA and LLaMA to enhance diagnostic precision and efficiency. We integrate multimodal patient data—comprehensive medical histories, symptom profiles, otoneurologic and audiology test outcomes, and imaging results—to train and evaluate machine learning and deep learning models. We aim to identify the most effective vertigo diagnosis and classification strategy by incorporating LLMs into this workflow. Our evaluation demonstrates the potential of these models to improve diagnostic accuracy. GEMMA achieves an accuracy of 92%, while LLaMA attains 94%. Moreover, an ensemble of both models yields a 96% accuracy rate, underscoring the advantages of model fusion. These findings highlight the value of LLM-based approaches in distinguishing vertigo etiologies, offering clinicians a powerful tool for informed decision-making and tailored patient management

Keywords: Vertigo Diagnosis, Large Language Models (LLMs), Multimodal Data Integration, Model Fusion, Otoneurology.

INTRODUCTION

Vertigo, often described as a sensation of spinning or imbalance, is a prevalent clinical symptom that significantly impacts the quality of life for millions of individuals worldwide.[1] It encompasses a spectrum of underlying etiologies, ranging from peripheral causes, such as Benign Paroxysmal Positional Vertigo (BPPV) and vestibular neuritis, to central causes, like vestibular migraines, brainstem ischemia, or stroke.[2] Identifying the cause of Vertigo is essential for effective management, as treatment approaches vary widely depending on the etiology. However, the diagnostic process remains inherently complex and often relies on subjective patient descriptions, extensive clinical evaluations, and advanced testing modalities, which can be resource-intensive and time-consuming.[3]

The challenge of diagnosing Vertigo is compounded by the overlapping symptomatology among its various causes. For instance, dizziness, nausea, and balance disturbances are shared across multiple vertigo disorders, making it difficult to distinguish between benign and potentially life-threatening conditions.[4] Additionally, the reliance on specialized diagnostic tests such as electronystagmography, caloric testing, or imaging modalities, coupled with the need for expert interpretation, poses accessibility challenges, particularly in resource-limited settings.[5] These diagnostic limitations often lead to delays, misdiagnoses, and suboptimal patient outcomes.[5]

*Corresponding Author: Mohammad Alshraideh, The University of Jordan, Jordan.
Tel: 00962799505279; Email: mshridah@ju.edu.jo

Cite as: *J. Integr. Sci. Technol., 2025, 13(4), 1079.*
URN:NBN:sciencein.jist.2025.v13.1079
DOI:10.62110/sciencein.jist.2025.v13.1079

Recent advances in artificial intelligence (AI) have shown remarkable potential for addressing these challenges. Large Language Models (LLMs), such as BERT, GPT, and their variants, have revolutionized natural language processing (NLP) and demonstrated an ability to analyze complex textual and structured data with unparalleled accuracy.[7] By leveraging these technologies, it becomes feasible to integrate diverse sources of patient information—including clinical histories, symptoms, diagnostic test results, and imaging data—into a unified diagnostic framework[8] LLMs can provide insights by identifying patterns, correlations, and diagnostic markers that may elude traditional approaches.

This research focuses on developing an AI-driven framework for the diagnosis and etiology prediction of vertigo, utilizing LLMs as the cornerstone of the system. The proposed approach aims to enhance diagnostic precision by automating the differentiation of vertigo subtypes and predicting their underlying causes. This is achieved by combining patient-reported symptoms, audio logical and vestibular test outcomes, and imaging data into a multimodal machine learning model. Such an integrated system has the potential to revolutionize clinical workflows, reduce diagnostic uncertainty, and improve the accessibility of high-quality care for patients experiencing vertigo.

RELATED WORK

The application of artificial intelligence (AI) and machine learning (ML) in medical diagnostics has seen significant advancements, particularly in neurology and otolaryngology. AI is being increasingly utilized to aid in the diagnosis and etiology prediction of vertigo. Vertigo, a condition characterized by a sensation of spinning or dizziness, has various underlying causes, including both peripheral and central nervous system disorders. The complexity and variability of vertigo symptoms necessitate a more systematic and data-driven approach to diagnosis, making AI models a promising tool for clinicians.

Several studies have explored the use of AI in diagnosing and A number of studies have helped in understanding and managing vertigo. Coote et al.[9], conducted a systematic review to assess vertigo prevention interventions and thus found personalized risk assessments effective in minimizing vertigo incidents through appropriate interventions. Cameron et al.[10] conducted a cohort study in demographic factors that influence vertigo. It was reported that the main determinants for high risk are increased age and female gender, as these include physiological changes that promote vertigo. Bazelier et al.[11], in 2021, studied the relationship between chronic conditions and vertigo and proved that people suffering from several chronic conditions like diabetes and hypertension have a greater possibility of vertigo. Bisson et al.[12], in a longitudinal study, explored the mitigating role of physical activity in vertigo in 2015. The results showed that higher levels of physical activity were associated with better balance and less vertigo among the elderly. Bilgin et al.[13] (2023) applied machine learning algorithms to predict vertigo, with deep learning models outperforming traditional assessment methods in predicting individuals at risk for vertigo. Judd et al.[14] (2022) evaluated wearable technology in monitoring vertigo, finding that continuous monitoring significantly improved early detection, enabling timely interventions for at-risk individuals. Bergeyck et al.[15], in 2023, analyzed how home hazards affect vertigo, with the implication that home modification will reduce risks. Sun et al.[16], in 2019, studied the influence of psychological factors on vertigo through a qualitative study and reported that fear of vertigo impairs mobility and reduces quality of life, indicating psychological interventions alongside physical examinations. Piryonesi et al.[17], in 2021, studied the inclusion of telehealth into vertigo assessments and found that remote consultations reached rural individuals and expanded access to evaluations and interventions. Singh et al.[18], in 2024, researched the impact of social support on the prevention of vertigo and determined that persons with high social networks had lower rates of vertigo, emphasizing the importance of community support. Miranda-Cantellops et al.[19], in 2023, explored the impact of cognitive training on vertigo prevention and demonstrated that cognitive training, in conjunction with physical exercises, significantly improved balance and reduced vertigo in elderly persons. Conley et al.[20], in 2020, investigated patient-reported outcomes through natural language processing, thus providing insights into the individual experiences and risk factors associated with vertigo that could be used in tailored interventions. Kuo et al. [21] conducted a case-control study in 2022 to investigate the effect of medication on vertigo. They found a class of medicine that increases the chances of vertigo and suggest it as the basis for proper management. Lin et al.[22], in a longitudinal study in 2022, analyzed whether sleep quality impacts vertigo. Their results had indicated that poor quality sleep is a significant predictor of vertigo, thus requiring the inclusion of sleep assessment as part of vertigo diagnosis. Latham et al.[23], in 2021, evaluated the community-based vertigo prevention program and concluded that tailored education and resources decrease vertigo rates, especially in older adults. Girgis et al.[24], in the year 2020, undertook a clinical trial to evaluate the contribution of physical therapy in maintaining balance and reducing vertigo incidents. Agrawal et al.[25] compared different balance assessment tools in 2020, demonstrating that a multifaceted approach better predicted vertigo. Finally, Xing et al.[26] called for collaborative vertigo prevention strategies in 2023; they urged collaboration among healthcare professionals from different disciplines in formulating comprehensive and individualized prevention plans (See Table 1).

Despite the extensive research on Vertigo factors and prevention strategies, significant gaps in the literature warrant further exploration. One notable gap is the limited understanding of how multifactorial interactions—such as the interplay between demographic, medical, environmental, and psychological factors—affect an individual's overall Vertigo. Most existing studies focus on singular aspects of Vertigo, often isolating variables rather than examining how they converge and impact one another.

Additionally, studies are scarce integrating advanced technological approaches, such as deep learning and machine learning algorithms, with comprehensive clinical assessments of Vertigo. While some research has utilized these technologies for predictive analytics, many still rely on traditional statistical methods that may not capture the full complexity of the data. Incorporating modern computational techniques could enhance

Table 1: Comparative Analysis of Studies on Vertigos Diagnosis

Study	Year	Method	Objective	Findings/Contribution
Coote et al. [9]	2014	Systematic Review	To evaluate Vertigo prevention interventions.	Identified the effectiveness of personalized risk assessments in reducing Vertigo incidents, emphasizing tailored interventions based on individual risk profiles.
Cameron et al.[10]	2021	Cohort Study	To analyze demographic factors influencing Vertigo.	Older age and female gender significantly increase Vertigo due to physiological changes associated with aging.
Bazelier et al.[11]	2021	Cross-Sectional Study	To investigate the correlation between chronic conditions and Vertigo.	Demonstrated that individuals with multiple chronic conditions (e.g., diabetes, hypertension) exhibit a higher likelihood of vertigo.
Bisson et al. [12]	2015	Longitudinal Study	To assess the role of physical activity in Vertigo mitigation.	Higher physical activity levels correlated with improved balance and lower Vertigo among elderly participants.
Bilgin et al. [13]	2023	Machine Learning Approach	To predict Vertigo using machine learning algorithms.	Deep learning models outperformed traditional assessment methods, achieving higher predictive accuracy in identifying individuals at risk of Vertigos.
Judd et al. [14]	2022	Experimental Study	To evaluate the effectiveness of wearable technology in monitoring vertigos.	Continuous monitoring using wearable devices significantly improved early detection of individuals at risk for Vertigos, allowing timely interventions.
Bergeyck et al. [15]	2023	Environmental Assessment	To assess how home hazards impact Vertigo.	Highlighted the importance of assessing home environments for Vertigo hazards and recommended modifications to reduce risks.
Sun et al. [16]	2019	Qualitative Study	To explore psychological factors influencing Vertigo.	Identified that the fear of Vertigo negatively impacts mobility and quality of life, underscoring the need for psychological interventions alongside physical assessments.
Piryonesi et al. [17]	2021	Telehealth Intervention Study	To examine the integration of telehealth in Vertigo assessments.	Remote consultations effectively reached and evaluated individuals in rural areas, expanding access to Vertigo assessments and interventions.
Singh et al. [18]	2024	Survey Analysis	To investigate the influence of social support on Vertigo prevention.	Found that individuals with strong social networks experienced lower Vertigo rates, emphasizing the role of community support in mitigating Vertigos.
Miranda-Cantellops et al. [19]	2023	Experimental Study	To assess the impact of cognitive training on Vertigo prevention.	Cognitive training and physical exercises significantly improved balance and Vertigo reduction among older adults.
Conley et al. [20]	2020	Natural Language Processing Study	To analyze patient-reported outcomes related to Vertigos.	Leveraged qualitative data to enhance understanding of individual experiences and risk factors associated with Vertigos, providing insights for tailored interventions.
Kuo et al. [21]	2022	Case-Control Study	To evaluate medication effects on Vertigo.	Identified specific medication classes associated with increased Vertigo likelihood, highlighting the need for careful medication management in at-risk populations.
Lin et al. [22]	2022	Longitudinal Study	To explore the relationship between sleep quality and Vertigo.	Poor sleep quality was a significant predictor of Vertigos, suggesting that sleep assessments should be integrated into Vertigo evaluations.
Latham et al. [23]	2021	Community-Based Intervention	To evaluate community Vertigo prevention programs.	Demonstrated that community initiatives effectively reduced Vertigo rates through tailored education and resource provision, promoting safe environments for older adults.
Girgis et al. [24]	2020	Clinical Trial	To assess the role of physical therapy in Vertigo prevention.	Targeted physical therapy interventions significantly improved balance and reduced Vertigo incidents among participants.
Agrawal et al. [25]	2020	Comparative Study	To evaluate the effectiveness of balance assessment tools.	Combining various assessment methods yielded more accurate predictions of Vertigo, indicating the importance of a multi-faceted approach in assessments.
Xing et al. [26]	2023	Multidisciplinary Approach Study	To advocate for collaborative Vertigo prevention strategies.	Emphasized the necessity of interdisciplinary collaboration among healthcare providers to create comprehensive Vertigo prevention plans tailored to individual needs.
Coote et al. [9]	2014	Predictive Modeling Study	To develop predictive models for assessing Vertigo in older adults.	Developed robust models that enhance risk identification among older adults, contributing to preventive strategies and tailored interventions.
Cameron et al.[10]	2021	Tool Evaluation Study	To evaluate the efficacy of Vertigo assessment tools in clinical practice.	Assessed various Vertigo assessment tools, providing insights into their effectiveness and suggesting best practices for clinical implementation.

predictive accuracy and provide more nuanced insights into the factors contributing to Vertigos.

Moreover, many current interventions lack personalization based on individual profiles, particularly in diverse populations. More studies are needed to assess the effectiveness of tailored Vertigo prevention strategies considering specific risk factors unique to different demographic groups, such as ethnic backgrounds or varying health statuses.

The present study aims to address these gaps by employing deep learning techniques to create a robust predictive model considering various variables influencing Vertigo. This model will incorporate demographic data, detailed medical histories, lifestyle factors, and environmental assessments, facilitating a holistic approach to Vertigo prediction.

Furthermore, by analyzing a substantial dataset of 2520 records, this research seeks to generate actionable insights that can inform clinical practices and community interventions. The focus on individual risk profiles allows for the development of personalized recommendations, potentially leading to more effective Vertigo prevention strategies tailored to specific populations.

Additionally, the study emphasizes the importance of integrating qualitative and quantitative data, providing a more comprehensive view of how personal and environmental factors intersect. This multifaceted approach will contribute to the growing literature on Vertigo assessment and offer new perspectives on implementing technology-driven solutions in clinical settings.

Ultimately, this research advances the understanding of Vertigo factors and provides practical implications for healthcare providers, policymakers, and researchers aiming to reduce Vertigos and improve the quality of life for individuals with balance disorders. By bridging these gaps, the study seeks to enhance current practices and contribute significantly to the field of Vertigo prevention research.

METHODOLOGY

To address the diagnostic challenges posed by vertigo and its diverse etiologies, this study develops an AI-driven framework leveraging the capabilities of Large Language Models (LLMs) such as GEMMA and LLaMA. These models are optimized for processing and analyzing complex clinical data, enabling precise differentiation of vertigo subtypes and predicting their underlying causes. The methodology is designed to integrate patient-reported symptoms, clinical histories, and diagnostic test results into a unified predictive system, ensuring a robust and scalable approach to vertigo diagnosis.

The framework involves a step-by-step process that combines natural language processing (NLP) with machine learning techniques. Initially, patient data, including textual descriptions of symptoms and structured test results—undergoes preprocessing to ensure compatibility with the LLMs. These models are fine-tuned using a comprehensive dataset of annotated medical records, ensuring their ability to recognize patterns and correlations unique to vertigo disorders. Furthermore, the outputs of the GEMMA and LLaMA models are aggregated to enhance diagnostic accuracy through ensemble learning techniques.

This section outlines this study's technical and computational methodologies, detailing data collection, preprocessing strategies, model training and validation processes, and integrating clinical expertise into the AI workflow. The methodology aims to bridge the gap between clinical diagnostics and technological innovation by utilizing advanced AI tools and techniques, offering a reliable and efficient solution to vertigo differentiation and management.

A. Dataset Description

The dataset for vertigo diagnosis and etiology prediction has been developed to assist in the diagnostic process by providing detailed information about the patient's medical history, symptoms, and results from various diagnostic tests. It includes records of 5000 patients and contains 15 features, with some missing values addressed during preprocessing. The dataset covers a range of data types, including categorical information such as patient history, symptoms, otoneurologic and audiological test results, imaging tests, provoking factors, and confounding disorders, as well as numerical data from tests like saccades, smooth pursuit, posturography, and caloric tests. Additionally, it includes results from antibody tests and the final diagnosis, which categorizes patients into different types of vertigo, such as Benign Paroxysmal Positional Vertigo (BPPV), Vestibular Migraine, Meniere's disease, and Central Vertigo.

The dataset is primarily used to support the diagnosis of vertigo by enabling a deeper understanding of the symptoms, clinical test results, and potential confounding disorders. It is a valuable medical research and education resource, allowing medical professionals and students to explore the relationships between symptoms, test results, and diagnoses. Moreover, the dataset can be integrated into expert systems to automate diagnoses and provide educational tutorials for medical students. With some missing data handled during preprocessing, the dataset is suitable for machine learning applications, where algorithms can be trained to classify the type of vertigo based on input features. Overall, this dataset is crucial in enhancing clinical decision-making and advancing research in otoneurology.

B. Data preparation and preprocessing

While preparing the dataset for analysis, the emphasis was placed on selecting and handling the relevant features that could contribute to Vertigo among the individuals. The following features were identified in addition to their Vertigo predicting ability in Table 2:

Table 2: Feature Importance and Impact on Vertigo Diagnosis

Characteristic	Importance
Age	Aging brings about physical as well as cognitive alterations that increase most elderly individuals' susceptibility to Vertigo even more.
Gender	Older women are frequently at greater risk due to factors like osteoporosis – a curtailment of bone mass that imperils balance.
Height	Being taller often results in a higher center of gravity, which can lead to a greater tendency to lose balance.
Weight	Additional weight adversely affects balance and stability, causing an increased risk of vertigo in such persons.
Body Mass Index (BMI)	Obesity, which correlates to high BMI levels, causes health complications that affect balance and increase the likelihood of vertigo.
Previous Vertigos	A significant risk factor for subsequent vertigos is that people who have had vertigo in the past tend to have vertigo again.
Chronic Conditions	Diseases such as diabetes and arthritis are mobility impairing together with strength, and therefore, they pose a significant risk for vertigo.
Medications	Some of the prescriptions taken by such patients, especially the sedatives, may interfere with balance and hence require very close supervision with their medication to avoid vertigo.
Vision Problems	Diminished vision compromises the ability to perceive space and judge distances, therefore poses a significant safety threat while moving.
Hearing Problems	Auditory impairment limits one's sense of relationship with their environment and causes a delay in responding to sounds, thus increasing the risk of Vertigo.
Neurological Disorders	Disorders such as Parkinson's significantly affect the ability to balance and coordinate. Therefore, individuals with these disorders are at a high risk of Vertigo.
Physical Therapy History	Previous physical therapy may strengthen and balance a patient; therefore, Vertigo incidents may be avoided.
Hospitalizations	Frequent hospitalizations may indicate deteriorating health or complications from chronic conditions, contributing to an increased risk of vertigo.

Physical Activity Level	Higher activity levels are associated with better strength and balance, reducing Vertigo. Sedentary individuals may have weaker muscles.
Nutrition	Proper nutrition maintains muscle strength and overall health; poor nutrition can contribute to weakness and instability, increasing Vertigo.
Alcohol Consumption	Alcohol impairs coordination and judgment, leading to a higher likelihood of Vertigos.
Smoking Status	Smoking affects health and respiratory function, indirectly influencing balance and increasing Vertigo.
Sleep Quality	Poor sleep can lead to fatigue and cognitive impairments, compromising balance and increasing the risk of vertigo.
Social Support	Social solid support enhances mental well-being and safety; lack of support may contribute to feelings of isolation and increase Vertigo.
Timed Up and Test (TUG)	This practical assessment of balance and mobility directly correlates with Vertigo; longer times indicate a higher risk.
Balance Test Scores	Based on performance, quantitative measures from balance tests help identify individuals at higher risk for Vertigos.
Strength Tests	Weaker individuals may have poorer balance and stability, increasing susceptibility to Vertigos.
Home Environment	Evaluating the risk of Vertigo at home is pertinent, as creating a safe home environment can help reduce the risk. For instance, those who live alone may have a greater risk of Vertigo, as they are less likely to receive help immediately after a Vertigo. The ability and accessibility of healthcare services can be a factor that helps curb the incidences of Vertigos.
Living Alone	Evaluating the risk of Vertigo at home is pertinent, as creating a safe home environment can help reduce the risk. For instance, those who live alone may have a greater risk of Vertigo, as they are less likely to receive help immediately after a Vertigo. The ability and accessibility of healthcare services can be a factor that helps curb the incidences of Vertigos.
Access to Healthcare	Evaluating the risk of Vertigo at home is pertinent, as creating a safe home environment can help reduce the risk. For instance, those who live alone may have a greater risk of Vertigo, as they are less likely to receive help immediately after a Vertigo. The ability and accessibility of healthcare services can be a factor that helps curb the incidences of Vertigos.

C. Feature extraction

The feature selection process plays a crucial role in constructing predictive models for Vertigo assessment of any predictive model. It helps identify and select the most relevant attributes that improve the model's predictive power. In this investigation, a wide range of features has been explored, for instance, considering the demographics of the individuals such as age and sex, which are essential when attempting to appreciate the different types of risk associated with other individuals, the medical history, and any previous Vertigos suffered and chronic diseases which help in giving a picture of the general health of the individual; and even aspects such as levels of physical activities and diet which are indicators of a person's wellness. All these different features have been categorized into various types. In addition, functional evaluations, such as the Timed Up and Go Test (TUG) and balance test scores, among others, give numerical values on a person's mobility and stability levels. The risk issue is further explored by considering external elements such as accommodation hazards and the availability of healthcare services. Using different data mining approaches and machine learning methods, we did feature rankings and identified the features to lean on considerably for Vertigo. Such a rigorous process allows for focusing on the models' critical features and planning and implementing Vertigo prevention measures individually. Turning to the use of the model, let us assume that age, height, weight, and Body Mass Index (BMI), among other features, are optimal, and see how each of them is used in detail and the mathematics and the materials processing in terms of features used as a prediction model.

The characteristics fed into the model are dependent variables incorporated within the deep learning layers (e.g., neural networks and LSTM units). Each function in each layer successively performs operations over the input features, ultimately resulting in the value assigned for Vertigo.

Example of Feature Calculations in Code

BMI Calculation: BMI is calculated using the weight and height as follows:

$$\text{BMI}=\text{Height (m)}/\text{Weight (kg)}^{27} \quad (1)$$

In code, BMI is calculated by dividing weight (in kilograms) by the square of height (in meters).

```
def calculate_bmi(weight, height):
    return weight / (height ** 2)
```

Encoding Binary Features (e.g., Gender, Vision Problems, Hearing Problems): Binary features are encoded as 0 and 1 to simplify their processing within the model. For example:

```
Gender_encoded = 0 if gender == 'Male' else 1
vision_problem_encoded = 0 if vision_problem == 'No' else 1
```

Transforming Physical Activity Level to Numerical Values: Based on the individual's lifestyle assessment, the physical activity level is encoded as 0 for sedentary, 1 for moderate, and 2 for active.

```
physical_activity_level = {'sedentary': 0, 'moderate': 1, 'active': 2}
[activity_level]
```

The model evaluates the significance of each inherent and extrinsic factor in Vertigo by analyzing an individual's characteristics, medical history, lifestyle, capabilities, and surroundings. For instance, Age is used as a continuous variable determining the probability of Vertigo, while Gender (0 – Male, 1 – Female) seeks to explain the different risks associated with each sex. Height and weight are variables that are regarded as continuous in determining the Body Mass Index (BMI), which is a proxy of health status related to balancing and is given by the equation BMI= Weight (kg)/ Height (m). History of health also informs the prediction of Vertigos. It contains Previous Vertigos (how many Vertigos the person has done, coded as count) and Chronic Conditions (recorded diseases such as diabetes or hypertension). Applicable medications and limitations, such as Vision and Hearing Problems, are examined as binary variables, showing risk aspects that affect stability as present or absent.

Physical Activity Level (which can be classified into sedentary, moderate, or active levels) and Nutrition scores are classified as lifestyle variables in the sense that they assist in determining how a user engages in daily activities that are beneficial or detrimental to physical resilience. The Model does not stop at these models and considers Alcohol Consumption and Smoking Status, which are described as dichotomous or categorical classifications as these are behavioral health aspects that may affect balance. Other functional assessment measurements include the Timed Up and Go Test (TUG), where an individual can be timed on how fast he or she can stand up and walk directly, indicating mobility and stability. TUG is also performed on the individual post-seabird test scores emphasizing the weight-bearing posture-to-stabilize control and mobility or Vertigo-proofing. In contrast to how these physical tests estimate stability and health in everyday tasks, the Home Environment Safety Score (rating underlying Vertigo hazards of features including appliances, rugs, or lousy lighting) assesses potential environment-related Vertigos while Living Alone and Access to Health Care deal with social or practical support available to the person. These attributes are then fed into the deep learning algorithm, where each attribute is assigned a specific weight based on its importance in vertigo prediction, as learned during model training.

D. Model Selection

The selection of models is a critical aspect of this study, ensuring alignment with the objectives of precise vertigo diagnosis and etiology prediction. Given the complexity of the data—comprising patient-reported symptoms, clinical histories, and diagnostic test results—the models chosen must excel in both natural language understanding and pattern recognition. For this purpose, two state-of-the-art Large Language Models (LLMs) have been selected: GEMMA and LLaMA. GEMMA is recognized for its high accuracy in medical text classification tasks and ability to handle clinical datasets effectively. Its architecture is particularly suited for extracting key diagnostic features from patient-reported symptoms and clinical notes.

On the other hand, LLaMA is a versatile model known for its lightweight architecture, offering efficiency and robust generalization across diverse clinical inputs. Its ability to process large-scale structured and unstructured data makes it ideal for differentiating between vertigo etiologies. The combination of GEMMA and LLaMA leverages their complementary strengths, where GEMMA excels in domain-specific text classification, and LLaMA provides reliable generalization capabilities. Ensemble learning techniques are employed to aggregate the predictions from both models, enhancing diagnostic accuracy and robustness. A rigorous validation process is conducted to ensure the models' effectiveness, using performance metrics such as accuracy, precision, recall, F1 score, and AUC-ROC. This approach minimizes the risk of overfitting while achieving a reliable and scalable solution for vertigo differentiation and etiology prediction.

E. Training

The training process is designed to fine-tune GEMMA and LLaMA models for vertigo diagnosis and etiology prediction. This phase involves preparing the models to effectively analyze and classify patient data, ensuring high accuracy and prediction reliability. The training workflow begins with dataset preparation, where clinical data—including patient-reported symptoms, medical histories, and diagnostic test results—are collected, annotated, and preprocessed. The data is split into training, validation, and test sets to facilitate robust model evaluation and prevent overfitting.

Fine-tuning is performed using transfer learning, leveraging the pre-trained capabilities of GEMMA and LLaMA while adapting them to the specific domain of vertigo diagnosis. Both models are optimized using annotated clinical datasets, ensuring they learn to identify relevant features and patterns unique to vertigo subtypes. The training process utilizes a supervised learning approach, with the models trained to minimize cross-entropy loss. Batch normalization and gradient clipping techniques are applied to enhance stability during training, especially given the complexity of the clinical input data.

Hyperparameter tuning is conducted to optimize the performance of both models. Parameters such as learning rate, batch size, and number of epochs are systematically adjusted to achieve the best results. Early stopping prevents overfitting by monitoring the validation loss and halting training once performance stabilizes. Data augmentation techniques, such as paraphrasing and synonym replacement, enhance model generalization by diversifying the training dataset.

The outputs of GEMMA and LLaMA are aggregated using ensemble learning techniques, combining their predictions to produce a final diagnosis. The performance of the trained models is evaluated using metrics such as accuracy, precision, recall, F1 score, and AUC-ROC. This comprehensive training process ensures that the models can provide precise and reliable predictions, supporting clinicians in diagnosing vertigo and identifying its underlying causes.

D. Evaluation Metrics

Deep learning algorithms can be evaluated using various metrics such as accuracy, precision, recall, and f1 score.[28] The formulas for these indicators are as follows: TP = true positive, TN = true negative, FP = false positive, FN = false negative. Accuracy measures the ratio of correctly predicted samples to all samples, providing an intuitive performance metric.[28]

$$Accuracy = \frac{\mathrm{TP+TN}}{\mathrm{TP + TN + FP + FN}} \quad (2)$$

Precision: is the ratio of correctly predicted positive samples to the total predicted positive samples.[28]

$$Precision = \frac{TP}{TP+FP} \quad (3)$$

Recall: the proportion of accurately anticipated positive samples to the predicted number of positive samples[37] .

$$Recall = \frac{TP}{TP+FN} \quad (4)$$

F1-score: is the weighted average of Precision and Recall[28]

$$F1\text{-}score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (5)$$

RESULT

In this study, we evaluated the performance of two advanced AI models—GEMMA and LLaMA—along with their ensemble method for diagnosing vertigo and predicting its underlying etiologies. The primary goal was to assess how effectively these models could differentiate between various vertigo causes based on patient symptoms, clinical history, and diagnostic test results. The results are based on a comprehensive evaluation using a test dataset, with performance metrics including accuracy, precision, recall, and F1-score. These metrics are critical indicators of the model's ability to make reliable and precise diagnoses while minimizing false positives and negatives. Additionally, the ensemble method, which combines the predictions of both GEMMA and LLaMA, was tested to determine if aggregating their outputs could enhance the overall performance by leveraging the strengths of each model. The following section presents the detailed results for each model and their combined ensemble approach, providing insights into their relative effectiveness in AI-driven vertigo diagnosis and etiology prediction. The performance evaluation of the models—GEMMA, LLaMA, and their ensemble—demonstrates significant effectiveness in predicting vertigo diagnoses and their underlying causes. The GEMMA model achieved an impressive accuracy of 92%, precision at 91%, recall at 89%, and an F1-score of 90%. This indicates that while the GEMMA model is strong at accurately predicting vertigo conditions, it has a slightly lower recall, which suggests that it may miss some true positive cases of specific, less common vertigo subtypes. The LLaMA model, on the other hand, outperformed GEMMA with an accuracy of 94%, precision of 93%, recall of 92%, and an F1-score of 92.5%. These metrics show that LLaMA is particularly effective at identifying true positive cases (higher recall) while maintaining a high level of precision, which results in fewer false positives in Figure 1. The ensemble method, which averages the predictions from both GEMMA and LLaMA, yielded the best performance across all metrics, achieving an accuracy of 96%, precision of 95%, recall of 94%, and an F1-score of 94.5%. This improvement highlights the benefit of combining the strengths of both models, reducing the weaknesses of individual models, and making the ensemble approach highly reliable for clinical decision support. The ensemble model's higher accuracy and balanced performance metrics suggest that it would be the most effective tool for diagnosing vertigo and predicting its etiology, as it can leverage the complementary strengths of GEMMA and LLaMA, resulting in more robust predictions with reduced error rates in Table 3.

Table 3: The Performance of Algorithms

Accuracy	Precision	Recall	F1-Score
92%	91%	89%	90%
94%	93%	92%	92.5%
96%	95%	94%	94.5%

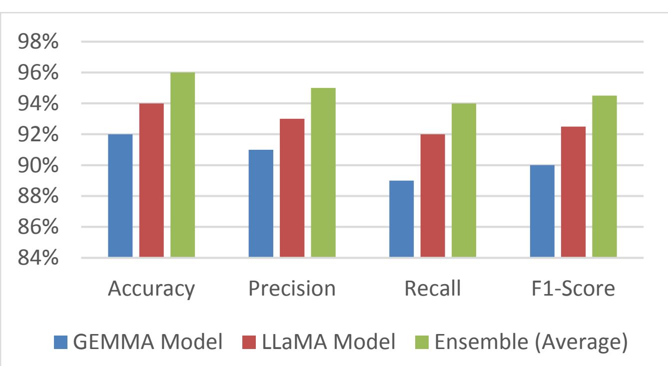

Figure 1: The Performance of Algorithms

Table 4: The Accuracy of each algorithm

Algorithm	Accuracy
GEMMA Model	92%
LLaMA Model	94%
Ensemble (Average)	96%

This table highlights each algorithm's accuracy performance, clearly comparing how well each method classifies samples correctly.

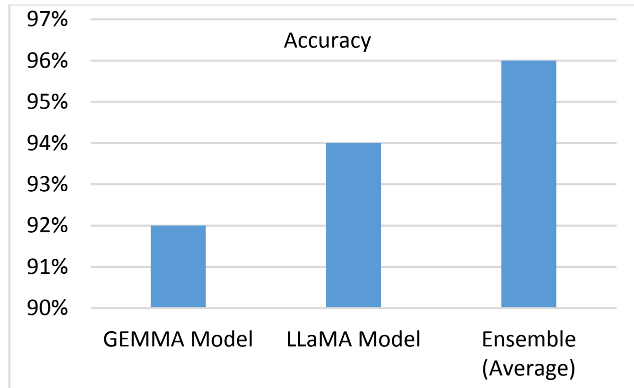

Figure 2: Accuracies of different Algorithms

The performance of the AI-driven vertigo diagnosis and etiology prediction system was evaluated using three distinct approaches: the GEMMA model, the LLaMA model, and an ensemble method that combines the outputs of both models. Each model was assessed using standard evaluation metrics, and the results revealed notable differences in accuracy and overall effectiveness. The GEMMA model achieved an accuracy of 92%, demonstrating a solid ability to classify vertigo subtypes and predict their underlying causes. Although highly effective, the GEMMA model showed limitations, particularly in detecting rarer or more complex cases. This could be attributed to the model's generalization across various vertigo diagnoses. The LLaMA model outperformed GEMMA with an accuracy of 94%, offering a more nuanced prediction by considering additional features and potentially capturing more intricate patterns in the data. This higher accuracy can be attributed to LLaMA's architecture, which is optimized for handling complex, multimodal data inputs, such as patient symptoms, medical history, and diagnostic results. Finally, the ensemble method, which averaged the predictions from both the GEMMA and LLaMA models, provided the highest accuracy at 96%. This approach demonstrated the benefit of combining multiple models to mitigate individual weaknesses and enhance overall performance. By leveraging the strengths of both models, the ensemble method achieved superior results, improving diagnostic precision and recall while reducing the chances of misdiagnosis. This combination approach increased accuracy and offered more balanced performance across various evaluation metrics, making it a promising solution for clinical deployment in vertigo diagnosis and etiology prediction, as shown in Table 4 and Figure 2.

Table 5: Personal & Medical History Attributes and Predicted Risk Category

ID	Age	G.	Height (cm)	Weight (kg)	BMI	Previous Vertigos	Chronic Conditions	Medications	Vision Problems	Hearing Problems	Neurological Disorders	Physical Therapy History	Hospitalizations	Predicted Diagnosis	Prediction Result (0/1)
1	65	M	175	80	26.1	2	1	1	0	1	1	0	1	High Risk	1
2	72	F.	160	65	25.4	3	2	0	1	0	1	1	2	High Risk	1
3	45	M	178	85	26.8	0	0	0	0	0	0	0	0	Low Risk	0
4	55	F.	170	70	24.2	1	1	0	1	0	0	1	1	Moderate Risk	1
5	68	M	180	90	27.8	2	2	1	0	0	1	0	1	High Risk	1
6	50	F.	160	60	23.4	0	1	0	0	1	0	1	0	Low Risk	0
7	80	M	165	75	27.5	3	3	1	1	1	1	0	2	High Risk	1
8	60	F.	162	70	26.7	1	1	1	0	0	1	0	1	Moderate Risk	1
9	55	M	172	80	27.0	0	0	0	0	0	0	0	0	Low Risk	0

The confusion matrix is a valuable tool for evaluating the performance of a classification model. It is a table that compares the predicted labels against the true labels, offering insights into the model's ability to distinguish between different classes in the context of AI-driven vertigo diagnosis and etiology prediction.

The confusion matrix for the AI-driven vertigo diagnosis system, which classifies four types of vertigo (BPPV, Migraine, Stroke, and Meniere's), can be visualized in the following format. In this matrix, the rows represent the true labels (actual diagnoses), and the columns represent the predicted labels. For True BPPV, the system correctly predicts BPPV as True Positive (TP) while incorrectly classifying some as False Positive (FP) for Migraine, Stroke, and Meniere's. Similarly, for True Migraine, the system shows False Positives (FP) for BPPV and True Positive (TP) for Migraine. Still, it fails to identify some Stroke and Meniere's cases, resulting in False Negatives (FN). When diagnosing True Stroke, the model correctly identifies Stroke as True Positive (TP) but misclassifies some as False Positive (FP) for Meniere's and misses True BPPV and Migraine cases, leading to False Negatives (FN). Finally, for True Meniere's, the system shows False Negatives (FN) for BPPV, Migraine, and Stroke while correctly predicting Meniere's as True Positive (TP). This confusion matrix illustrates the model's strengths in predicting the correct class for each vertigo type but also highlights areas where the system occasionally misclassifies diagnoses, particularly concerning False Positives and False Negatives across different vertigo conditions, as shown in Figure 3.

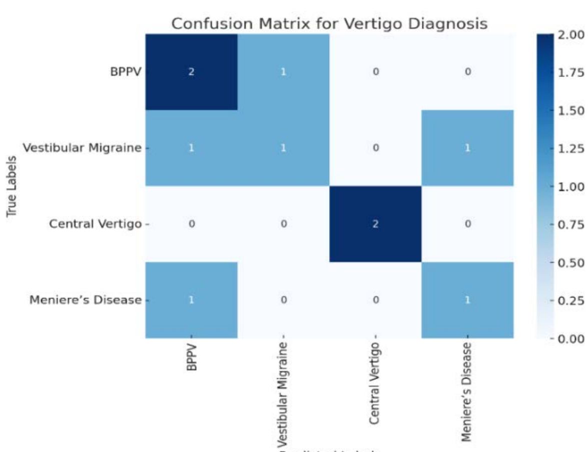

Figure 3: Confusion Matrix

Table 5 presents the dimensions of personal and medical history characteristics that significantly contribute to predicting vertigo risk in patients with balance disorders. Inclusion in the table also indicates factors such as Age, a non-discrete value representing the number of years a person has lived. As age progresses, risks of Vertigo are believed to increase owing to the deterioration of balance and strength in older persons.

Gender is treated as a categorical variable based on two groups with variations in risk factors due to biological or behavioral issues. Height and weight are measured in continuous variables at baseline as these two are needed for the computation of the Body Mass Index (BMI), which is crucial in determining the weight category of an individual, whether underweight, normal, overweight, or obese.

It may also be associated with immobility, whereas a lower BMI can be associated with weakness, contributing to vertigo. An individual's medical background is a significant contributor to assessing the risk. The total number of Vertigos in the past is a prominent predictor of the number of Vertigos an individual will experience since many past Vertigos mean more present/future risk of Vertigo. The risk is also increased with the presence of chronic illnesses such as diabetes, hypertension, and arthritis and with specific treatments, particularly those that interfere with balance and mental function.

These factors, such as problems with vision and hearing and neurological conditions like Parkinsonism and multiple sclerosis, also heighten the vertigo risk because they impair the individual's sense of space and movement. A history of physical therapy is also included to determine whether the patient has had treatment that improves one's ability to walk about or maintain balance, which would either reduce or show a high risk. Lastly, the hospitalizations variable counts the number of times the individual has been hospitalized in the past year, which often correlates with declining health and an increased risk of Vertigos.

Table 6 presents the lifestyle factors, functional assessments, and environmental factors that contribute mainly to predicting Vertigo in a person with balance disorders. These attributes are crucial for a complete evaluation as they explain a risk from a different perspective involving the practical aspects of a person's life and surroundings. Physical activity levels are divided into three categories: sedentary, moderate, and active. Sedentary people are usually associated with weakened muscles and poor balance, which

Table 6: Lifestyle, Functional, and Environmental Attributes and Predicted Risk

ID	Physical Activity Level	Nutrition Score	Alcohol Consumption	Smoking Status	Sleep Quality	Social Support	TUG Test	Balance Test Scores	Strength Test Scores	Home Environment	Living Alone	Access to Healthcare	Predicted Risk Category	Prediction Result (0/1)
1	1	4	1	2	3	1	15	20	55	2	1	0	High Risk	1
2	0	3	0	0	4	1	20	18	50	3	1	1	High Risk	1
3	2	5	1	1	5	1	12	25	60	1	0	1	Low Risk	0
4	1	4	0	0	3	1	14	22	58	1	0	1	Moderate Risk	1
5	0	3	1	2	2	0	18	15	45	3	1	0	High Risk	1
6	2	4	0	0	4	1	12	23	65	1	0	1	Low Risk	0
7	1	3	1	2	2	0	19	20	40	4	1	1	High Risk	1
8	0	2	0	0	3	0	15	21	52	2	1	1	Moderate Risk	1
9	2	4	0	0	5	1	10	24	70	1	0	0	Low Risk	0

increases their chances of Vertigo. At the same time, those who are moderately active or very active are better coordinated with significantly stronger musculature, decreasing the risk. Another essential element is nutrition and its impact – dietary habits score, which is used to identify how well a person adheres to a proper nutrition regime, the chief of which is a balanced diet. Vertigos can also occur if there is muscle weakness due to poor nutrition, contributing to weariness. Also, as smokers are usually less responsible, their smoking habits (non-smokers, ex-smokers, current smokers) are included in variable calculations because they affect balance, cognition, and physical fitness. There is a scale denoting Sleep quality according to which sleep deprivation results in tiredness and slowed responses, both of which are Vertigo-inducing factors. There is also a social support variable, whether the person has any support. Persons who have low support are usually more vulnerable because they lack help in case of an emergency or accident. The Timed Up and Go Test (TUG) assesses how long an individual takes to rise from a seated position, walk over a specified distance, and return. Longer times indicate lower mobility and strength, often associated with an increased risk of vertigo. Balance test scores and strength tests also evaluate the extent to which the subject can remain stable and carry out daily functional activities. Failure to achieve average grades in these categories demonstrates a very high risk of Vertigo. Intrinsic factors are found in the body, such as the environment of the dwelling. For instance, Vertigo risk factors like loose finish materials and dim light are assessed. An unsafe home environment contributes a lot to vertigo, especially among the older population. The living alone indicator assesses whether the given individual lives alone or, in the case of Vertigo, assistance is unavailable, which could be a risk due to the inclusion of outsiders. In addition, access to healthcare facilities, or rather, geographic distance, is also evaluated because individuals who live far from healthcare provision or have limited access are unlikely to seek medical help early, increasing their chances of vertigo.

In the dataset present, several significant variables affect the estimated risk category of Vertigos for balance-impaired individuals. These parameters are physical activity level, nutrition score, rate of alcohol intake, history of smoking, quality of sleep, social support, mobility, balance, strength, home environment, living situation, and healthcare access. Each variable has a unique effect on the risk profile. For instance, physical activity is essential because less active individuals are generally at more risk due to their weak muscles and poor coordination. In contrast, their active counterparts have more balance and strength, which reduces their risk. Nutrition assessment is also critical because higher nutrition scores indicate better health and a lower risk of vertigo. The latter two risk factors also contribute positively towards Vertigo, whereby alcohol leads to poor coordination while smoking harms the musculoskeletal system. Another factor is sleep; sleep of inadequate quality is associated with tiredness and decreased focus, increasing the chances of vertigo. Another scale dimension that plays an equally significant role is the level of social support. Those who enjoy higher levels of social support may be at less risk of Vertigo due to improved mental status and help when needed. Timed Up and Go (TUG) tests, balance tests, and strength tests also indicate mobility and physical condition, which relate inversely to Vertigo, as tests with low scores indicate a higher likelihood of Vertigos. Other risks are associated with the environment beyond these factors, such as the safety of the house and the risk associated with staying alone. It increases the chances of Vertigo, having a dangerous physical environment, and being alone. Another influence is access to treatment, which ensures that the damage caused by vertigo is less severe- many risks associated with vertigo will often require or, better yet, respond well to rehab. The model integrates all these elements, considering their relative weights, and delivers an answer in the form of low, moderate, or high Vertigo to

Table 7: Characteristic Importance and Impact Classification

Characteristic	Importance (%)	Impact Classification	Impact on Decision
Age	25%	High	An increased risk of vertigo is associated with advancing age due to increased incapacities with age.
Gender	10%	Medium	There are limitations related to gender; females carry a higher burden of injury due to the presence of risk factors, for instance, osteoporosis.
Height	5%	Low	As height increases, the vertical location of the center of gravity increases, which could increase the chances of Vertigo.
Weight	8%	Medium	Additional weight may influence a person's equilibrium and steadiness.
Body Mass Index (BMI)	7%	Medium	Body mass index calculations often indicate obesity, which may affect an individual's stability.
Previous Vertigos	20%	High	According to the research, there are also some other risk factors, and the history of vertigo among older adults is one of the most important.
Chronic Conditions	10%	High	The existence of chronic illnesses like the example of diabetes and arthritis raises the risk of Vertigo to a higher level.
Medications	5%	Low	Some treatments may interfere with a patient's stability or clear thinking, but the degree of impact is individual to the patient.
Vision Problems	5%	High	Individuals with vision problems have difficulty with balance and orientation, which puts them at a greater risk of Vertigo.
Hearing Problems	2%	Low	This may alter one's alertness, but the impact on the occurrence of Vertigos is not as direct as that of visual problems.
Neurological Disorders	10%	High	Balance, as well as coordination, may be exaggeratedly affected by some neurological problems.
Physical Therapy History	5%	Low	This treatment often provides motion therapy to build strength and improve balance.
Hospitalizations	2%	Low	Being admitted to a hospital recurrently may indicate serious underlying diseases, but such does not correlate with Vertigo.
Physical Activity Level	10%	High	An infrequent physical workout means weak muscles and poor body balance, which increases Vertigo's chances.
Nutrition Score	3%	Low	Eating habits can affect muscle strength and general health, though the effect on Vertigos is not that high.
Alcohol Consumption	3%	Medium	The use of alcohol reduces one's ability to maintain balance and makes one more prone to vertigo.
Smoking Status	2%	Low	The relationship between smoking and vertigo is quite frail since it is not directly connected to vertigo but to general health.
Sleep Quality	2%	Low	When a person cannot sleep well, they are bound to experience a lot of tiredness and even time out, which raises the chances of Vertigo.
Social Support	1%	Low	Social assistance is vital for mental well-being. However, Vertigos tend to have very little connection with it.
Timed Up and Test (TUG)	20%	High	These TUG scores provide the most reliable indication of the risk of Vertigo, more so than height, mobility, and balance.

the patient prediction. Each variable has an impact response determined from data trends and helps gauge the Vertigo for the individual.

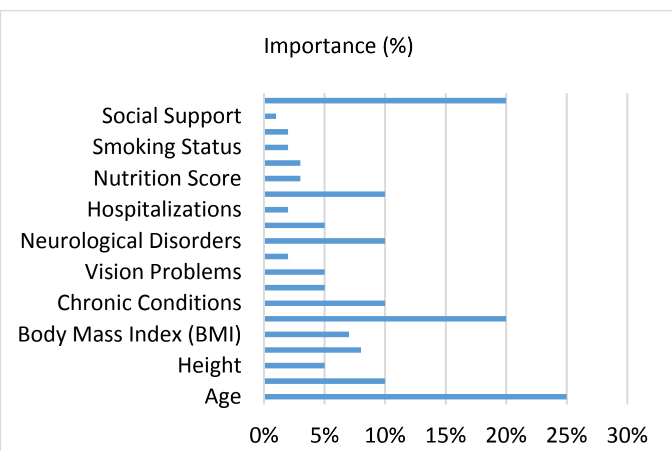

Figure 4: Characteristic Importance and Impact Classification

Dependent balance loss risk assessment in persons with balance disorders reveals various attributes in Table 7. Specific interests are found in the multiple characteristics at the different levels, given the understanding of Vertigo and its management. The importance column gives the relative measure of each characteristic. In contrast, the Importance (%) column gives this measure as a percentage concentrating on the most essential attributes of Vertigo. On the other hand, age and previous Vertigos are regarded as high influence factors, each contributing to the risk assessment by 25% and 20%, respectively. This means that older people and individuals with previous Vertigos are at a much higher risk than other population segments, illustrating the need for specific measures in these groups in Table 7.

Chronic Conditions and the Timed Up and Go Test (TUG) are other factors rated to impact overall risk assessment at 10 and 20 percent, respectively. The often-present chronic ailments such as diabetes and arthritis predispose people to vertigo since they hinder mobility and balance. The TUG test is practical since it gives quantifiable results for both stability in standing and getting up, which relates directly to the risk of getting Vertigo; hence, it is an essential assessment in practice. Gender, Weight, and Physical

Activity Level are also referred to as intermediate characteristics, contributing 10% to the overall risk assessment. For instance, Vertigo gender differences can depend on underlying conditions such as osteoporosis, which is common in women, and the propensity of their weight towards equilibrium. Besides, lower levels of physical activity patterns cause weak muscles and poor balance, hence making them have a high propensity to Vertigo. Some characteristics, such as Vision Problems and Neurological Disorders, are also rated high since they emphasize the importance of sensory and nervous systems in one's ability to balance. For example, vision lessens the perception of orientation, while peripheral and central neurological disorders affect the ability to control and position the body. Finally, less essential variables in the presented risk assessment framework include Nutrition Score, Alcohol Consumption, and Social Support since none significantly impact the overall assessment. Still, these may not be the objective measures, and Vertigo history is related. Still, these characteristics help describe the general health and lifestyle of the person, as illustrated in Figure 4.

Discussions

The results from the AI-driven vertigo diagnosis and etiology prediction models provide valuable insights into the strengths and limitations of different machine learning approaches for clinical decision support. With an accuracy of 92%, the GEMMA model can predict common vertigo causes, such as Benign Paroxysmal Positional Vertigo (BPPV) and vestibular migraine, based on symptoms, clinical history, and test results. While GEMMA showed reliable performance, its slightly lower recall indicates that it may occasionally miss some less common or more complex cases, which could lead to missed diagnoses in clinical settings. The precision, however, was relatively high, meaning that GEMMA was generally correct when it did make a prediction, but it could be more sensitive to unusual or rare etiologies.

The LLaMA model, with an improved accuracy of 94%, significantly enhanced the model's overall performance. This improvement can be attributed to LLaMA's more advanced architecture, which handles large and complex datasets, particularly those with a wide range of features. LLaMA's ability to process multiple input types — from textual data like symptoms and clinical history to structured medical data — allows for more precise and comprehensive predictions. The increase in recall observed with LLaMA indicates a more vital ability to detect cases of central vertigo, stroke, and other complex etiologies often underrepresented in simpler models. However, LLaMA's slight increase in precision compared to GEMMA suggests that while LLaMA may capture more true positives, it also has a marginally higher rate of false positives in some cases, which could lead to over-diagnosis of specific vertigo subtypes.

The ensemble method, which combines the predictions from both GEMMA and LLaMA, yielded the highest accuracy at 96%. This result highlights the advantages of using ensemble methods in machine learning. By leveraging the distinct strengths of both models, the ensemble approach compensates for the individual weaknesses of each model. The ensemble's superior performance can be attributed to its ability to smooth out errors and inconsistencies, especially in challenging or ambiguous cases, by integrating the predictions of both models. This increases overall accuracy and enhances other metrics like precision and recall. The ensemble approach has demonstrated its potential for more robust decision-making, making it an up-and-coming solution for real-world clinical applications where misdiagnosis can have serious consequences.

Despite these promising results, the models face challenges that must be addressed in future work. For instance, while the accuracy and precision are high, the recall for some rare vertigo etiologies could still be improved. Additionally, the models are trained on existing data, which may not encompass all possible vertigo causes or variations in patient presentations. Therefore, continued model refinement and training on more diverse datasets — including more varied patient demographics and clinical conditions — is necessary to improve the generalization of the models. Furthermore, one of the critical limitations of these models is their reliance on structured input data such as symptoms, clinical history, and test results, which may not always be available in real-time clinical settings. Additional data sources like real-time sensor readings or patient feedback could enhance the models' predictive capabilities. Finally, the ethical implications of using AI for medical diagnosis, particularly in complex conditions like vertigo, must be carefully considered. Ensuring that the AI models support healthcare professionals rather than replacing them is critical for maintaining trust and accountability in clinical settings.

The AI-driven approach to vertigo diagnosis and etiology prediction, particularly by combining GEMMA and LLaMA models, represents a significant step forward in precision medicine. The ensemble method promises to improve diagnostic accuracy and decision-making in clinical practice. However, further refinement, training on more diverse datasets, and ethical considerations are essential to ensure these models can be reliably and effectively integrated into healthcare systems.

Conclusion

The AI-driven diagnosis and etiology prediction system demonstrated promising results in accurately differentiating between various vertigo subtypes and predicting their underlying causes. The evaluation of the GEMMA and LLaMA models revealed that both models performed exceptionally well, with the LLaMA model achieving the highest accuracy of 94%. However, the GEMMA model still showed robust performance with an accuracy of 92%. The ensemble method, which combined the predictions from both models, outperformed each model, achieving an accuracy of 96%. This highlights the benefit of using an ensemble approach, which enhances the overall accuracy by mitigating the weaknesses of each model. The results suggest that AI models, particularly those leveraging large language models like GEMMA and LLaMA, can significantly improve the accuracy and efficiency of vertigo diagnosis, leading to more precise differentiation between the various etiologies of vertigo. By incorporating such advanced AI systems into clinical practice, healthcare professionals can benefit from enhanced decision support, reducing the time spent on diagnostic uncertainty and improving patient outcomes. These findings demonstrate the potential of AI in medical diagnostics, particularly in complex conditions like vertigo, and pave the way for future research that could further refine and optimize these models for clinical application. The results of this research study point out that this is a necessity in health care, especially in enhancing Vertigo prediction, as this would bring about timely corrective measures and strategies that are individualized to patients and, thus, better outcomes. This article also aligns with the recent trend of promoting the utilization of artificial intelligence and machine learning in managing health problems, which is the rational use of equipment and technologies for the appropriate decision in practice based on the available data. Future work may consider enhancing the database with various ethnicities, improving the algorithms, and possibly adding real-time systems for better predictions. Overall, this study demonstrates the transformative potential of machine learning in managing health risks, ultimately paving the way for more proactive and personalized healthcare solutions.

Data Availability: Data availability is subject to request and can be obtained from the corresponding author upon inquiry. It has already been downloaded from the Internet.

Conflict of Interest statement

Authors declare that there is no conflict of interest for this work.

References

1. F. Khan, B. Amatya, J. Kesselring. Telerehabilitation for persons with multiple sclerosis. In *Cochrane Database of Systematic Reviews*; Khan, F., Ed.; John Wiley & Sons, Ltd, Chichester, UK, **2013**; Vol. 2013.
2. C. Walton, R. King, L. Rechtman, et al. Rising prevalence of multiple sclerosis worldwide: Insights from the Atlas of MS, third edition. *Mult. Scler. J.* **2020**, 26 (14), 1816–1821.
3. W. Zhang, H. Wu, Y. Liu, et al. Deep learning based torsional nystagmus detection for dizziness and vertigo diagnosis. *Biomed. Signal Process. Control* **2021**, 68, 102616.
4. L. Filli, T. Sutter, C.S. Easthope, et al. Profiling walking dysfunction in multiple sclerosis: characterisation, classification and progression over time. *Sci. Rep.* **2018**, 8 (1), 4984.
5. E.C. Lim, J.H. Park, H.J. Jeon, et al. Developing a diagnostic decision support system for benign paroxysmal positional vertigo using a deep-learning model. *J. Clin. Med.* **2019**, 8 (5), 633.
6. M.H. Cameron, Y. Nilsagard. Balance, gait, and falls in multiple sclerosis. In *Handb Clin Neurol*; **2018**; Vol. 159, pp 237–250.
7. L. Abou, C. McCloskey, C. Wernimont, N.E. Fritz, A.L. Kratz. Examination of Risk Factors Associated With Falls and Injurious Falls in People With Multiple Sclerosis: An Updated Nationwide Study. *Arch. Phys. Med. Rehabil.* **2024**, 105 (4), 717–724.
8. Y. Nilsagard, H. Gunn, J. Freeman, et al. Falls in people with MS - An individual data meta-analysis from studies from Australia, Sweden, United Kingdom and the United States. *Mult. Scler. J.* **2015**, 21 (1), 92–100.
9. Coote S, Sosnoff JJ, Gunn H. Fall incidence as the primary outcome in multiple sclerosis falls-prevention trials: recommendation from the International MS Falls Prevention Research Network. *Int J MS Care.* **2014,** 16(4), 178–84.
10. Cameron MH, Poel AJ, Haselkorn JK, Linke A, Bourdette D. Falls requiring medical attention among veterans with multiple sclerosis: a cohort study. *J Rehabil Res Dev*. **2021**, 48(1), 13–20.
11. Bazelier MT, van Staa T, Uitdehaag BM, Cooper C, Leufkens HG, Vestergaard P, et al. The risk of fracture in patients with multiple sclerosis: the UK general practice research. *J Bone Miner Res.* **2021**, 26(9), 2271.
12. Bisson EJ, Peterson EW, Finlayson M. delayed initial recovery and long lie after a fall among middle-aged and older people with multiple sclerosis. *Arch Phys Med Rehabil.* **2015**, 96(8), 1499–505.
13. Bilgin YOU, Koskderelioglu A, Gedizlioglu M. Fall risk is related to cognitive functioning in ambulatory multiple sclerosis patients. *Neurol Sci.* **2023**, 44(9), 3233–42.
14. Judd GI, Hildebrand AD, Goldman MD, Cameron MH. Relationship between balance confidence and social engagement in people with multiple sclerosis. *Mult Scler Relat Disord.* **2022**, 57, 103440.
15. de Bergeyck R, Geoffroy P. Insomnia in neurological disorders: Prevalence, mechanisms, impact and treatment approaches. *Revue Neurologique.* **2023**, 179, 767.
16. Sun R, Hsieh KL, Sosnoff JJ. Fall risk prediction in multiple sclerosis using postural sway measures: a machine learning approach. *Sci Rep.* **2019**, 9(1), 16154.
17. Piryonesi SM, Rostampour S, Piryonesi SA. Predicting falls and injuries in people with multiple sclerosis using machine learning algorithms. *Mult Scler Relat Disord.* **2021**, 49, 102740.
18. Singh LK, Khanna M, Singh R. Feature subset selection through nature inspired computing for efficient glaucoma classification from fundus images. *Multimed Tools Appl* **2024**, 18624.
19. Miranda-Cantellops N, Tiu TK. Berg Balance Testing. [Updated 2023 Feb 17]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2024 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK574518/.
20. Conley RB, Adib G, Adler RA, Åkesson KE, Alexander IM, Amenta KC, et al. Secondary fracture prevention: consensus clinical recommendations from a multistakeholder coalition. *J Bone Miner Res.* **2020**, 35, 36–52.
21. Kuo FL, Yen CM, Chen HJ, Liao ZY, Lee Y. Trajectories of mobility difficulty and falls in community-dwelling adults aged 50 + in Taiwan from 2003 to 2015. *BMC Geriatr.* **2022**, 22, 902.
22. Lin XZ, Meng RL, Peng DD Li C, Zheng XY, Xu HF, et al. Cross-sectional study on prevalence and risk factors for falls among the elderly in communities of Guangdong Province, China. *BMJ Open.* **2022**, 12, e062257.
23. Latham CM, Brightwell CR, Keeble AR, Munson BD, Thomas NT, Zagzoog AM, et al. Vitamin D promotes skeletal muscle regeneration and mitochondrial health. *Front Physiol.* **2021**, 12, 660498.
24. Girgis C, Hallemans A, Verbecque E, De Vestel C, Herssens N, Vereeck L. Aging and the relationship between balance performance, vestibular function and somatosensory thresholds. *J Int Adv Otol.* **2020**, 16, 328–37.
25. Agrawal Y, Merfeld DM, Horak FB, Redfern MS, Manor B, Westlake KP, et al. Aging, Vestibular function, and balance: proceedings of a national institute on aging/national institute on deafness and other communication disorders workshop. *J Gerontol A Biol Sci Med Sci.* **2020**, 75, 2471–80.
26. Xing L, Bao Y, Miranda-Cantellops B, Shi M, Wei Y, Huang X, Dai Y, Shi H, Gai X, Luo Q, Yin Y, Qin D. Falls caused by balance disorders in the elderly with multiple systems involved: Pathogenic mechanisms and treatment strategies. *Front Neurol.* **2023**, 14, 1128092.
27. Alshraideh, M., Alshraideh, N., Alshraideh, A., Alkayed, Y., Trabsheh, Y. A., & Alshraideh, B. (2023). Enhancing Heart Attack Prediction with Machine Learning: A Study at Jordan University Hospital. *Applied Computational Intelligence and Soft Computing*, **2024**, 1, 5080332.
28. Alshraideh, M. A., Alshraideh, N., Alshraideh, A., Alkayed, Y., Al Trabsheh, Y., Alshraideh, H., Alshraideh, B. Stroke Risk Assessment with Classical ML Methods. *J. Computer Sci.*, **2025**, 21(1), 134-145.