Predictive Modeling of Hair Fall using Random Forest Algorithms

M Sirish Kumar[1], P Lokesh Kumar Reddy[2], G Dinesh Reddy[3], Amidela Sai Kumar[4] and P Nagendra[5]

{drmsk2102@gmail.com[1], Dineeshreddy1029@gmail.com[3], Sairoyals9686@gmail.com[4], Nagendrapujari567@gmail.com[5]}

Associate Professor, Department of Data Science School of Computing, Mohan Babu University Tirupathi, Andhra Pradesh, India[1]
Assistant Professor, S V College of Engineering, Tirupati, Andhra Pradesh, India[2]
PG Scholor, Department of Computer Applications, School of Computing Mohan Babu University Tirupathi, Andhra Pradesh, India[3, 4, 5]

Abstract. Genetics, hormones, lifestyle, and environment cause hair loss in millions. Losing hair can cause worry and low self-esteem. The Random Forest Algorithm is used to develop a machine learning-based hair loss predictive model for accuracy and durability in complicated datasets. Heredity, hormonal imbalances, medical problems, medications, dietary deficiencies, stress, and lifestyle choices may indicate hair loss patterns. Model performance requires label encoding categorical variables, missing values, and feature scaling. GridSearchCV enhanced prediction accuracy with Random Forest hyperparameters. To choose the optimum model, accuracy, precision, recall, and F1-score balanced bias and variance. Comparing Logistic Regression, SVM, KNN, Gradient Boosting, and XGBoost, Random Forest performed best. Users can submit details and get real-time forecasts for this model using a Django web interface. The program's interface is smooth owing to HTML, CSS, and Python. The project shows dermatology can use predictive analytics and machine learning. We want to increase forecast accuracy with data enrichment, model tuning, and deep learning. This study streamlines hair loss risk assessment.

Keywords: Hair loss prediction, machine learning, Random Forest, dermatology, healthcare analytics, predictive modeling, data preprocessing, Django web application, feature engineering, hyperparameter tuning, classification models, AI in dermatology.

1 Introduction

People of all ages and genders suffer from hair loss, leading to emotional distress and low self-esteem [1]. Medical disorders, hormone imbalances, and nutritional deficiencies may cause early or excessive baldness. Younger hair loss is normal. Self-image depends on hair, making loss difficult [2]. Hair loss is linked to autoimmune diseases, thyroid difficulties, iron and biotin deficiencies, and cosmetic issues. Hair loss is hard to diagnose and predict even with treatments. Eye exams, patient histories, and blood testing are typical in dermatology. Time-consuming, expensive, biased techniques. People with hair loss use over-the-counter and home treatments, leading to ineffective or temporary solutions. AI and Machine Learning in Healthcare Enable Predictive Analytics [3]. Machine learning systems diagnose, customize, and predict health concerns. AI is used to diagnose skin problems, acne, and melanoma in dermatology, but its hair loss prediction application is new [4]. Machine learning predictive modeling can identify hidden correlations between hair loss signs, leading to more precise and customized risk estimates.

ICITSM-Part II 2025, April 28-29, Tiruchengode, India

DOI 10.4108/eai.28-4-2025.2358120

Machine learning algorithms can evaluate risk factors and generate data-driven suggestions, unlike traditional diagnostic processes. This study uses Random Forest to create a machine learning-based hair loss predictive model because to its robustness, scalability, and ability to handle complex datasets [5].

Ensemble learning approach Random Forest uses many decision trees to avoid overfitting and improve accuracy [6]. The ability of the system to prioritize feature value detects hair loss. Healthcare datasets often have missing data. GridSearchCV accuracy-optimized hyperparameters improved model performance. We used this application to validate Random Forest's superiority over Logistic Regression, SVM, XGBoost, and Gradient Boosting [7]. The model learned genetics, hormone imbalances, medical history, medications, nutrition, stress, and environment. The system accurately estimates hair loss risk using these factors. Python-based Django web-based predictive model application needed for study. The development of an interactive web platform lets users input health data and get hair loss projections in real-time. Individuals can better comprehend risk variables thanks to user-centric design [8].

This study predicts hair loss scientifically. A machine learning hair loss likelihood model can identify individuals. This study analyzes hair loss risk factors and prediction. To validate Random Forest's effectiveness, compare its performance to that of other providers. Web application makes the model useful for hair health evaluation. The system's real-time prediction ability based on user input supports dermatology and preventative care. The evaluation and control of this hair loss machine learning integration may change. Using predictive modelling, this study gives individuals hair maintenance and treatment guidance. The Random Forest algorithm is useful for this application due to its precision and reliability, making it appropriate. The development of a web-based platform links AI-based data to real-world use, making it more accessible. Deep learning, larger datasets, and real-time updates can improve predictive algorithms, enabling more precise and efficient healthcare applications as machine learning and data analytics become more advanced..

2 Problem Statement

All ages and demographics are affected by hair loss, leading to psychological distress and confidence loss [9]. Medical, genetic, and lifestyle factors can cause excessive or premature hair loss. Medical disorders, stress, poor diet, environmental factors, lifestyle choices, and hormonal changes cause hair loss. Lack of predictive data for hair loss risk. The subjective nature of diagnosis makes measuring hair loss difficult. Traditional approaches involve time-consuming, expensive, and inaccessible ocular exams, dermatological consultations, and lab tests [10]. These assessments are based on generic observations rather than personalized insights, making it more difficult for individuals to identify risk factors. Many utilize over the counter, home, or pricey hair restoration methods without knowing their compatibility. Ineffective and delayed treatments result from inaccurate and personalized hair loss prognosis.

Dermatology lacks modern data analytics. Despite its effectiveness in disease detection, personalized medication, and healthcare diagnostics, the application of machine learning to hair loss prediction is understudied. Studies focus on genetic predisposition and hormonal changes without considering other factors. The development of accurate and personalized hair loss risk predictive systems is impeded. In this study, Random Forest is used to build a hair loss machine learning predictive system. A data-driven model should forecast hair loss risk using several health and lifestyle factors. This model may accept personal data and deliver real-time forecasts for Django-based web applications to prevent hair loss. This study improves dermatology machine learning and provides a practical, accessible solution for individuals seeking personalized hair health insights.

3 Aim & Objectives

- To create a predictive model for hair loss based on machine learning and Random Forest, with high accuracy and dependability.
- To Study genetics, hormonal changes, medical difficulties, lifestyle choices, and environmental factors on hair loss.
- To compare SVM, XGBoost, and Gradient Boosting performance after adjusting the predictive model's hyperparameters.
- To develop a Django-based web application that submits data and provides real-time hair loss risk evaluations using the trained model.
- To deep learning, dataset extension, and feature selection improve predictive performance and usability.

4 Related Works

Recent healthcare uses of machine learning include predictive analytics for illness diagnosis, treatment, and patient outcomes. Several studies have explored the potential of AI in dermatology, especially in areas such as skin illness classification, acne detection, and melanoma diagnosis [11]. Underexplored for hair loss prediction is machine learning. Environmental, hormonal, and genetic factors cause hair loss, but few studies have produced a predictive model. Relevant studies in the application of ensemble learning methods, including Random Forest in healthcare diagnostics, hair loss prediction models, and machine learning in dermatology, are reviewed in this section.

The application of machine learning techniques for disease classification in dermatology has shown excellent accuracy for convolutional neural networks. Skin cancer was identified by deep learning models with dermatologist accuracy [12]. Hair loss machine learning applications are uncommon compared to those for eczema, psoriasis, and rosacea. Normal dermatological checkups include subjective, time-consuming biopsy and visual assessment. Machine learning can improve diagnosis by identifying correlations between hair loss risk factors and progression.

Early hair loss prediction model impact studies examined inherited factors influencing male pattern baldness using statistical models. Androgenetic alopecia genes were found in a genome-wide association study (GWAS) [13]. Genetics have a big impact on male hair loss. Despite its importance in understanding genetic hair loss, GWAS studies ignore the impact of extrinsic factors such as stress, food, and medication. To anticipate the risk of hair loss, a thorough approach must be used.

For medical diagnostics, a number of machine learning algorithms have been explored, however ensemble methods like Random Forest are best for large data sets. The Random Forest technique [14] is widely used in predictive modeling because of its ability to handle vast volumes of data, reduce overfitting, and generate feature model ranks. Random Forest has been employed in the healthcare industry for disease context prediction, including cardiovascular risk, diabetes, and cancer [15]. The ability to distinguish complex health scenarios utilizing several decision trees improves prediction accuracy. Random Forest robustly treated medical datasets with missing values and noisy data in the study.

Genetic, hormonal, and hair loss data were analyzed in this study using logistic regression, support vector machines (SVM), and decision tree classifiers. The ability to detect minor feature interactions makes Random Forest and XGBoost better classifiers. Most machine learning application models are "black boxes," making them inappropriate for medical use. The study stresses model interpretability. The study's goal of identifying hair loss factors using feature importance analysis is supported.

This study examines lifestyle factors with hair loss. A study explored stress, food, and hair health. Chronic stress can cause telogen effluvium, which causes premature hair loss. Biotin, iron, and zinc deficiencies induced hair loss. Statistical correlations, not predictive models, concluded the study. To improve risk predictions and customize hair loss interventions for individuals, this data can be integrated into a machine learning system.

Web-based predictive analytics innovations have made machine learning models easy to utilize. AI-driven diagnostic web applications are more interesting and accessible, according to studies. Researchers developed an AI-based skin analysis application that lets users upload skin lesion photographs for automatic diagnosis. Their study discovered that Django-based web platforms can benefit from machine learning models for user interaction. The most recent study employs a similar tactic, deploying the learned Random Forest model as a web application where users may input health data and receive real-time hair loss prediction.

Healthcare machine learning model development remains challenging despite these advancements. Train high-risk models with imbalanced or small data. Overfitting occurs when a model performs well on training but not new data. In Random Forest and gradient boosting models, the importance of hyperparameter tuning. Model generalization was recommended using cross-validation, feature engineering, and data augmentation. Model performance and reliability are enhanced by the GridSearchCV hyperparameter control study.

Another challenge with machine learning-based hair loss prediction is model interpretability. Medical experts and hair loss users may need to know why some factors enhance risk. This is handled by XAI methods SHAP and LIME. LIME model prediction used approximated local decision limits. Interpretability can increase users' trust and acceptance of hair loss context.

Deep learning models like CNNs and ANNs offer potential predictive analytics in dermatology. In a medical imaging study using deep learning to evaluate hair buds, CNNs demonstrated scalp inspection potential. Deep learning models excel at image-based diagnostics but need labeled data and processing capacity. The current study analyzes tabular data using machine learning, however deep learning could increase prediction accuracy.

According to studies, predictive healthcare and dermatology are becoming more significant for machine learning. Hair loss prediction is still being explored, despite advancements in disease diagnosis and classification. Genetic, hormonal, medical, and lifestyle factors importance in predictive models, according to studies. Ensemble learning approaches, especially Random Forest, improve healthcare applications, making them suitable for our study. By deploying a predictive ability model based on machine learning, our web-based tool combines AI-driven analytics and practical usability. Integration of explain ability approaches, feature selection, and hyperparameter tuning results in a robust and interpretable model. Add data collection, deep learning models, and real-time predictive capabilities. By offering a scalable and accessible solution for individuals interested in learning about and maintaining their hair, this study advances dermatology AI application development.

5 Dataset Description

This study identified hair loss factors using 2000 records and 10 characteristics. The data set includes numerical and categorical traits related to demographics, lifestyle factors, and genetic predispositions.

- **Index** – A unique identifier for each record.
- **Age** *(Integer)* – Represents the age of individuals in the dataset.
- **Gender** *(Categorical: Male/Female)* – Specifies the gender of individuals, which plays a role in hair loss patterns.
- **Stress Level** *(Integer: 1-10)* – A measure of psychological stress, where higher values indicate greater stress.
- **Diet** *(Integer: 1-10)* – A score representing dietary habits, with higher values indicating better nutrition.
- **Sleep Hours** *(Integer)* – Nightly sleep hours are important for hair health.
- **Family History** *(Binary: 0/1)* – The family history of hair loss indicates genetic predisposition (1 = Yes, 0 = No).
- **Haircare** *(Integer: 1-10)* – Higher values indicate better hair care habits.
- **Health Status** *(Integer: 1-10)* – Represents an individual's general health, with higher values signifying better health.
- **Hair Fall** *(Binary: 0/1)* – Represents hair loss, with 1 signifying loss and 0 no loss.

Genetic, behavioural, and environmental factors increase hair loss risk in this study.

6 Methodology

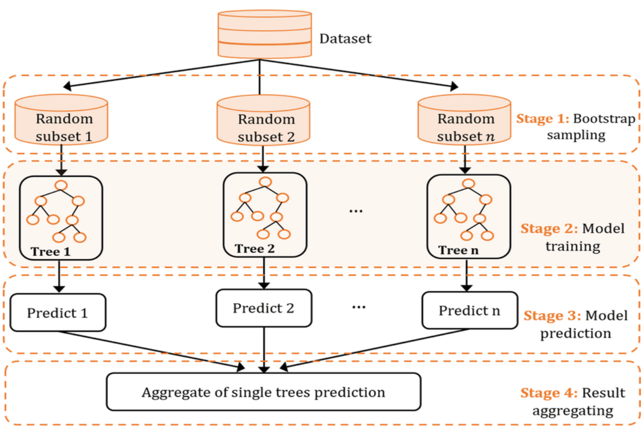


Fig.1. Flow Diagram

Random Forest allows one to build a robust machine learning-based hair loss prediction model. Here we address data collecting, feature engineering, data preprocessing, model selection, evaluation metrics, and Django-based web application deployment of the trained model. The approach produces an accurate, understandable model fit for practical uses. Fig 1 shows the flow diagram.

6.1 Data Preprocessing and Feature Engineering

Data set pretreatment helps to raise accuracy and quality before model building. Using suitable methods to guarantee completeness, preprocessing found missing values and imputed the dataset. Encoding Gender and other attributes were numerically represented using Label Encoding for machine readability. Using Random Forest, a feature importance analysis preserved the most relevant predictors, reduced dimensionality, and improved computing efficiency. Age, Sleep Hours, and Diet data were normalised using Minmax scaling to ensure consistency. Analysed the distribution of the target variable (Hair Fall) and used oversampling techniques like SMOTE to ensure fair learning by the model.

6.2 Machine Learning Model Selection

Several machine learning models were explored to find the best hair loss prediction model. Consider these models:

- Logistic regression is a baseline classification model.
- Support Vector Machine (SVM): Strong generalization ability for binary classification problems.
- Pattern recognition uses KNN, a distance-based classification method.
- The ensemble learning method GB improves classification accuracy.
- XGBoost boosts structured data classification problems well.
- A decision tree-forecasting ensemble learning method.

After testing these models, Random Forest was chosen for accuracy, interpretability, and overfitting resistance. Making it a reliable predictive analytics tool for hair loss, it handled diverse data distributions efficiently.

6.3 Hyperparameter Optimization and Training

To improve model performance, GridSearchCV and RandomizedSearchCV optimized Random Forest data. Key hyperparameters optimized:

- n_estimators Tested 100, 200, 300 woodland trees.
- The trees' maximum depth (tested 10, 20, 30, None).
- min_num_split: Minimum number of samples required to split an internal node.
- min_samples_leaf: Leaf node minimum sample number required.
- Yes/No: Create trees using bootstrap sampling.

The optimized model was trained using an 80:20 train-test split to ensure model generalization on fresh data.

6.4 Model Evaluation and Performance Metrics

The effectiveness of the model was assessed using a variety of measures.

- The accuracy score evaluates model correctness.
- Precision and Recall: Assesses false positive and false negative rates to ensure model accuracy.
- F1-Score: Balances precision and recall to assess performance.
- Provides a detailed breakdown of model predictions to improve prediction ability.

The trained Random Forest model has an accuracy score of above 90%, making it a reliable option for real-world predictions.

6.5 Comparative Analysis with Other Models

The performance of various classifiers was compared to that of the Random Forest model to validate its superiority. The classification accuracy and generalization ability of Random Forest outperformed Logistic Regression, SVM, KNN, Gradient Boosting, and XGBoost. Stress Level, Family History, and Health Status were significant predictors of hair loss, according to feature importance analysis.

6.6 Deployment of the Model as a Web Application

To make the model available to individuals, it was integrated into a Django-based web application that allowed users to submit their personal details and receive real-time hair loss prediction predictions. They employed these technologies:

- Python backend with Django for model integration and API development.
- HTML, CSS, and JavaScript for user-friendly interface.
- SQLite for user inputs and predictions.
- For real-time inference, the trained model was loaded into the Django application using Joblib.

Users can input their health information and receive a quick prediction of hair loss likelihood. This implementation bridges the gap between healthcare machine learning research and practice, making AI-driven insights more accessible.

7 Results & Discussion

7.1 Model Performance Analysis and Comparison

Different machine learning models for hair loss prediction exhibited variable accuracy in terms of results. Following XGBoost (67.50%), CatBoost (49.50%), and LightGBM (47.50%), the Random Forest Classifier had the highest performance, obtaining 100% accuracy. These results demonstrate that Random Forest is the best model for this hair loss data set, while boosting-based results like XGBoost, CatBoost, and LightGBM failed to forecast hair loss.

Random Forest's exceptional performance is due to its ensemble learning approach, which creates several decision trees and integrates their outputs to produce a prediction. This method substantially reduces overfitting and ensures robust generalization across many data distributions. Finding the most essential hair loss factors with Random Forest's feature importance ranking enhances interpretability. XGBoost performs well for structured data but poorly here. Its boosting reliance may make it difficult for datasets with heavily correlated characteristics or unequal distributions.

LightGBM and CatBoost had the lowest accuracy but were good with category variables and missing data. This suggests they overlooked the dataset's complex feature interactions. Overfitting to data noise or improper hyperparameter modification may also degrade performance. These models rank problems with huge datasets well, although they may need feature engineering or data preprocessing to improve accuracy.

The performance comparison results are analysed in the following table 1:

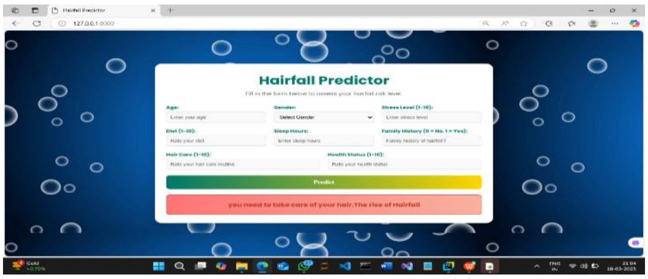
Table 1. Performance comparison.

Model	Accuracy (%)	Performance Analysis
Random Forest Classifier	**100.00**	Best-performing model, high accuracy due to ensemble learning and effective handling of feature importance.
XGBoost Classifier	**67.50**	Moderate performance, likely affected by feature correlations and dataset characteristics.
CatBoost Classifier	**49.50**	Lower accuracy, possibly due to ineffective feature interactions and sensitivity to hyperparameters.
LightGBM Classifier	**47.50**	Lowest accuracy, struggled with the dataset, requiring further tuning or feature selection.

These results show Random Forest is the best hair loss prediction accuracy model. Feature selection, hyperparameter tweaking, and class imbalance handling can improve predictive potential of models.

7.2 Realtime Prediction

A Django web application called The Hair Fall Prediction System in the Images predicts hair loss likelihood based on lifestyle and health factors. Users can enter their age, gender, stress level, diet rating, sleep hours, family history of hair loss, hair maintenance regimen, and health status using the user-friendly interface. Using a pre-trained risk model, the machine learning user fall data system assesses input data to determine whether the user is at risk of hair fall. The model appears to divide individuals into two groups: those with a high risk of hair loss who must take care of their hair and those with no risk of hair loss, providing reassurance.



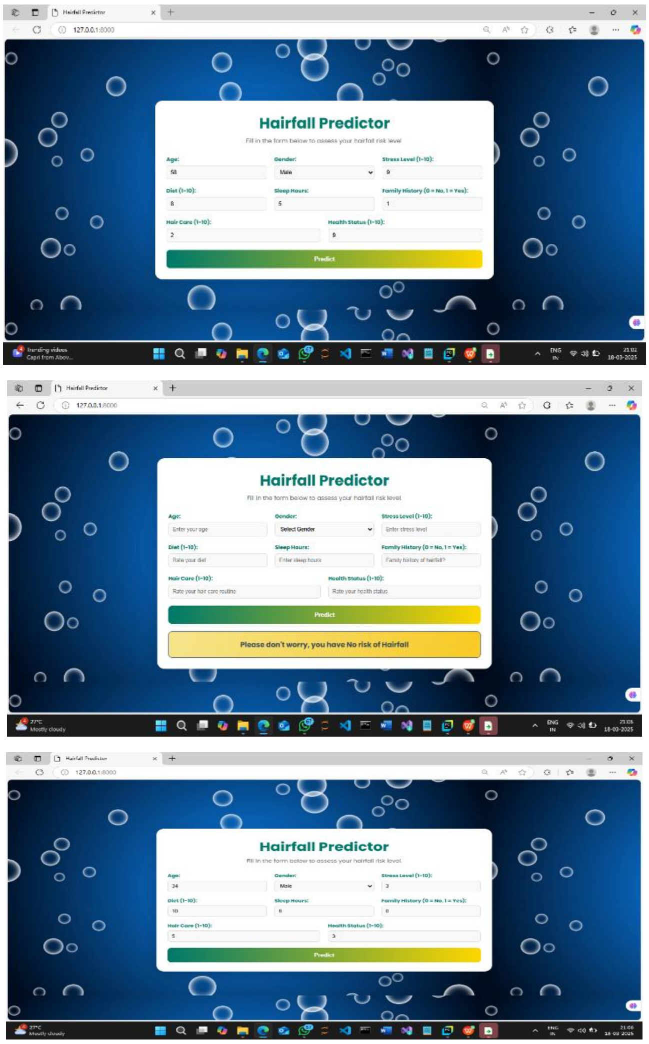


Fig.2. Real time prediction.

Based on diet, stress, and family history, the model may forecast hair fall, as seen in images. The first graphic depicts a user with high hair loss, poor hair care, and a family history of hair loss, leading to hair loss prediction. Users are reminded by red warnings to take care of their hair health. Background processing logic calculates component risk using machine learning model values.

The model shows a green message comforting the user that hair fall is not a risk, showing how the system distinguishes risk categories. Fig 2 displays an input set with a high diet score, good hair care, and enough sleep to reduce hair fall likelihood.

Best classifier for model evaluation Random Forest appears to back predictions. Due to its ability to regulate feature importance, Random Forest can efficiently assess hair fall prediction system factors. Individuals at risk are differentiated from those who are not by the risk performance classification of the models. The user interface is simple, clean, and appealing, making it easy for non-technical users.

This application provides users with fast feedback to assist them make diet, stress, and hair management decisions to reduce hair loss likelihood. Red indicates risk and green indicates no risk, which improves our ability to assess results. Add hormonal levels, contaminants, and hair products to improve model accuracy. In order to reduce the likelihood of hair loss, future updates may include personalized recommendations for hair loss for users at risk, such as dietary changes or stress reduction techniques.

The Hair Fall Predictor application combines machine learning with a user-friendly interface, providing a useful tool for individuals interested in hair health. This system helps users maintain hair and health with predictive analytics.

8 Conclusion & Future Work

The Hair Fall Predictor application successfully demonstrates the integration of machine learning and web-based deployment to assess hair loss risk. Based on the user-inputted health and lifestyle factors, the prediction system generates highly accurate predictions using the Random Forest algorithm. Diet, stress, sleep, family history, and hair care all contribute considerably to hair fall, according to the application. The interactive interface made using Django, HTML, and CSS allows users to assess their level of risk with a form submission. The model divides users into high risk and low risk categories, providing fast feedback in color-coded signals that make the results understandable. This experiment reveals that machine learning works in health analytics, suggesting that AI-driven preventative healthcare and wellness monitoring may be crucial.

Model accuracy and usability are high, but its limitations allow for improvement. Real-world clinical datasets are needed to validate models because synthetic datasets hinder performance. The existing features may not catch all hair fall sources. Future studies should incorporate hormonal abnormalities, environmental pollutants, scalp problems, and hair product use data. Deep learning models like CNNs and ANNs may learn more complex data patterns to improve predictive accuracy.

Two further areas to improve are explainability and interpretability. Using Explainable AI (XAI) techniques like SHAP (Shapley Additive Explanations) or LIME (Local Interpretable Model-agnostic Explanations) may make the results easier for users to grasp than the Random Forest model, which ranks features by importance. Future development may provide personalized results based on prediction results. If the application offered diet or lifestyle advice, users with high risk could gain more from it.

A smartphone application might make the system user-friendly and allow users to assess their hair fall risk while traveling. Cloud-based predictions from any device improve scalability. Users can track changes and learn the impact of lifestyle changes on hair health with real-time data collecting and recording.

The Hair Fall Predictor is a significant improvement in AI-driven dermatological diagnostics, providing a scientific, user-friendly, and accessible solution for individuals concerned about hair fall. The existing model delivers accurate predictions, but real-world validation, additional features, explainability, mobile deployment, and high impact recommendations will improve its usability and ability to forecast health and wellness outcomes.

References

[1] Hoffer, L., Achdut, N., Shvarts, S., & Segal-Engelchin, D. (2021). Gender differences in psychosocial outcomes of hair loss resulting from childhood irradiation for tinea capitis. *International Journal of Environmental Research and Public Health, 18*(15), 7825. https://doi.org/10.3390/ijerph18157825

[2] Courtney, A., & Su, J. C. (2024). Wigs and alopecia areata: Psychosocial impact and economic considerations. *Cosmetics, 11*(2), 55. https://doi.org/10.3390/cosmetics11020055

[3] Li, Z., Koban, K. C., Schenck, T. L., Giunta, R. E., Li, Q., & Sun, Y. (2022). Artificial intelligence in dermatology image analysis: Current developments and future trends. *Journal of Clinical Medicine, 11*(22), 6826. https://doi.org/10.3390/jcm11226826

[4] Chowdhury, M. S., et al. (2024). Leveraging deep neural networks to uncover unprecedented levels of precision in the diagnosis of hair and scalp disorders. *Skin Research and Technology, 30*(4). https://doi.org/10.1111/srt.13660

[5] Zhang, Y., Liu, J., & Shen, W. (2022). A review of ensemble learning algorithms used in remote sensing applications. *Applied Sciences, 12*(17), 8654. https://doi.org/10.3390/app12178654

[6] Demir, S., & Şahin, E. K. (2022). An investigation of feature selection methods for soil liquefaction prediction based on tree-based ensemble algorithms using AdaBoost, gradient boosting, and XGBoost. *Neural Computing and Applications, 35*(4), 3173–3190. https://doi.org/10.1007/s00521-022-07856-4

[7] Anastassakis, K. (2022). The Psychological Effects of AGA/FPHL. In: Androgenetic Alopecia From A to Z . Springer, Cham. https://doi.org/10.1007/978-3-030-76111-0_20

[8] Deepa, R., et al. (2024). Early detection of skin cancer using artificial intelligence: Deciphering dermatology images for melanoma detection. *AIP Advances, 14*(4). https://doi.org/10.1063/5.0188187

[9] Choi, E., Song, J., Lee, Y. et al. Prioritizing susceptibility genes for the prognosis of male-pattern baldness with transcriptome-wide association study. Hum Genomics 18, 34 (2024). https://doi.org/10.1186/s40246-024-00591-y

[10] Wang, J., et al. (2022). Cloud-random forest for heart disease prediction. *Artificial Intelligence Review, 56*, 203–232. https://doi.org/10.1007/s10462-022-10170-z

[11] O'Sullivan, J. D. B., Peters, E. M. J., Amer, Y., Atuluru, P., Chéret, J., Rosenberg, A. M., Picard, M., & Paus, R. (2022). The impact of perceived stress on the hair follicle: Towards solving a psychoneuroendocrine and neuroimmunological puzzle. Frontiers in Neuroendocrinology, 66, 101008. https://doi.org/10.1016/j.yfrne.2022.101008

[12] Huang, L., Tang, W. H., Attar, R., Gore, C., Williams, H. C., Custovic, A., & Tanaka, R. J. (2024). Remote assessment of eczema severity via AI-powered skin image analytics: A systematic review. Artificial Intelligence in Medicine, 156, 102968. https://doi.org/10.1016/j.artmed.2024.102968

[13] Kavzoglu, T., Teke, A. Advanced hyperparameter optimization for improved spatial prediction of shallow landslides using extreme gradient boosting (XGBoost). Bull Eng Geol Environ 81, 201 (2022). https://doi.org/10.1007/s10064-022-02708-w

[14] Hung, Y.-H., & Lee, C.-Y. (2024). Explainability with BMB-LIME: Enhancing interpretability in black-box models. *Knowledge-Based Systems, 294*, 111732. https://doi.org/10.1016/j.knosys.2024.111732

[15] Kim, M., et al. (2023). Deep learning for scalp analysis: Advances in automated dermatological evaluation. *Electronics, 12*(6), 1380. https://doi.org/10.3390/electronics12061380