

Detection of Stress based Hair fall and Prevention using Machine learning based KNN model

¹ G Pavan Kumar, ² A. N. L Kumar,

¹ MCA Student, Dept. Of MCA, Swarnandhra College of Engineering and Technology, Seetharampuram,
Narsapur, Andhra Pradesh 534280,

pavankumargandham9@gmail.com

² Associate Professor, Dept. Of MCA, Swarnandhra College of Engineering and Technology, Seetharampuram,
Narsapur, Andhra Pradesh 534280,

Abstract: Many things can have an effect on someone's mood, causing hair loss. Due to different factors which include robust personnel, excessive workload and overwork, employees inside the IT sector are, amongst different things, more harassed. Depression, tension, somatisation, and interest deficit disorder are simply some of the mental fitness problems that pressure can reason and be fatal. Therefore, it is vital to apprehend people's anxiety inside the early stages so that the appropriate remedies may be used to manipulate and decrease tension. A lot of research has been performed with the aid of predicting strain. The improvement of the skin, hair is the principle splendour of the character's face. The consequences of some learning algorithms, like KNN, are higher. Other synthetic intelligence strategies including ML algorithms may be used to diagnose diseases.

KEY WORDS- Machine learning, K-Nearest Neighbour Algorithm, Hair fall detection, Stress, Pressure.

I. INTRODUCTION

Hair at the scalp regulates body temperature and maintains brain temperature. A healthful character has 100,000 hairs on their head and generally loses 50 to one hundred hairs inside the afternoon. Seasonal regulations, horrible weight loss diets, micronutrient deficiencies, genetic resistance, and detrimental consequences of prescription

drugs are all elements that put stress on our day by day lifestyles. If the principle causes of hair loss are secondary, others might also journey. Hair transplants and prescription medicinal drugs are critical in some instances. Some ailments require antibiotic remedy due to a possible bacterial or fungal problem. Other uncommon conditions that lead to hair loss encompass alopecia, follicles and psoriasis.

Unlike alopecia, every day hair loss outcomes in the advent of coin-shaped areas of baldness and protects the complete scalp. Regular hair loss because of alopecia can be because of a selection of sicknesses. Hair may be misplaced” is a hair loss plan. An autoimmune sickness known as alopecia aerate causes immoderate hair loss which could then cover the whole scalp and cause hair loss. Millions of human beings around this quarter are in problem. Therefore it is able to be critical to observe new processes within the kind and evaluation of alopecia aerate. Different illnesses and situations can be classified and anticipated efficaciously via the purchase of astute knowledge within the tool.

II LITERATURE REVIEW

[1] "Deep Learning based Detection of Hair Loss Levels from Face Pictures", 2019 Halim Benhabiles, Karim Hammoudi, Ziheng Yang, Feryal Windal, Mahmoud Melkemi, Fadi Dornaika, and Ignacio Arganda-Carreras (IPTA)

Hair loss is a phenomenon regarded to have an impact on humans' morale and self-self warranty. Often, the attention of the phenomenon and the opportunities of treatment are late. This paper investigates deep analyzing strategies for detecting hairs loss ranges

thru the use of men from face pics. In this context, a specific education dataset has been organized with face pictures having several degrees of baldness. Moreover, notwithstanding the low visibility of hairs in such pics, an equal approach is proposed for robotically classifying facial pictures with appreciate to pattern type tables of male baldness from the clinical place. Experimental results display the capability and the performance for medical, safety and industrial packages.

[2] Survey-primarily based definitely Machine Learning techniques to diagnosis of hair fall illness in Bangladeshi Community," 2022 13th Farhana Khatun, Moshfiqur Rahman Ajmain, Sharun Akter Khushbu, Nushrat Jahan Ria, and Sheak Rashed Haider Noori (ICCCNT)

Hair symbolizes the splendour of women and men. Everyone is jealous of our hair. We lose our hair at a younger age due to some errors or routine things that we do. Many men and women across the world be afflicted by hair loss and the form of girls be afflicted by modifications in three hundred and sixty-five days. Genetically, dandruff, hypersensitive reaction and distortions are the main reasons of hair loss. We are doing this examine survey to help human beings. This appearance represents things. First, we study the number of reasons why hair loss is a cause

for concern. Another problem is that we teach our dataset with machine gaining knowledge of algorithms to decide accuracy. The gadget mastering technology has evolved faster than studies survey datasets. SVM, Logistic Regression, Naive Bayes, Decision Tree, Random Forest, K-Nearest Neighbour and XG Boost algorithms for complete performance assessment are well known. Experimental effects confirmed that the XG Boost had exquisite day by day performance, with an accuracy of 90. Sixty two%.

III System Analysis

Existing System

The result: the general evaluation isn't always executed in a timely way, which worsens the severity of the infection. To prevent lifestyles-threatening diseases along with cancers and tumours, network-based neural answers are used in diverse industries, including fitness and fitness informatics [12]. An overall of 150 images had been received from special locations and have been pre-processed to lessen mistakes rates thru extinction, averaging, evaluating, and enhancing image statistics. This equipment assists professionals and patients at the same time as presenting superior know-how of early symptoms and signs and symptoms. The three major forms of hair loss and scalp-related

problems that we did are expecting in this overview had been alopecia, psoriasis, and follicles. The venture, however, proved tough due to the paucity of studies on the situation, the shortage of a appropriate data set and the diploma of variability of pictures spread on the internet.

Disadvantages of existing machine:

The method of trichoscopy, which includes revealing the traits of hair loss thru scalp imaging, has been changed from that prescribed; but, the predictive impact become now not genuinely glaring while the use of snapshots.

Mechanically, it became new. Alopecia diagnosis promotes the usage of a combination of vision computer systems in photographic processing strategies.

Depending on the scalp scenario, the direct-acting synthetic paints trade depending on what the gadget uses to categorise the scalp pix.

The class accuracy of eighty-5 percentage modified with that obtained. In the unique examination,

Proposed System:

In this look at, the volume of hair loss is expected the usage of numerous influential traits. ML became used for this cause. Overall, 60% of the obtained information changed into used for instructional purposes, 20% for validation and the closing 20% for the neural community

verifying the special training algorithms had been used because of this. The effects of imposing those algorithms are blended. Neural networks had been proven to be effective at predicting hair loss. Thorny is a versatile and light-weight integrated improvement surroundings.

Advantages of current machine:

Unlike the present trichoscopy approach wherein the consequences of prediction were not virtually seen in scalp pixy, the proposed machine likely gives progressed predictive visibility. This might be because of better picture processing and laptop vision strategies that beautify the identity and visualization of hair loss traits.

The proposed system represents a singular approach to the diagnosis of alopecia aerate by integrating pc vision and photo processing strategies. This innovative technique can probably cause greater correct and efficient analysis as compared to traditional strategies.

IV EXPLORATORY DATA ANALYSIS

DATA ANALYSIS SETUP

DATA ANALYSIS SETUP

```
[3] import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

[4] import pandas as pd
data = pd.read_csv('/content/LUKE_H-1.csv')

[5] print(data)
```

	date	hair_loss	stay_up_late	pressure_level	coffee_consumed	\
0	01-01-2021	Few	2	Low	0	0
1	02-01-2021	Few	0	Low	0	0
2	03-01-2021	Medium	3	Low	1	1
3	04-01-2021	Few	2	Low	0	0
4	05-01-2021	Few	2	Low	0	0
...
395	31-01-2022	Medium	1	Low	1	1
396	01-02-2022	Few	1	Low	0	0
397	02-02-2022	Medium	1	Low	1	1
398	03-02-2022	Medium	0	Low	1	1
399	04-02-2022	Few	1	Low	0	0

DATA DESCRIBE

```
[6] data.describe()
```

	stay_up_late	coffee_consumed	brain_working_duration	hair_grease	libido
count	400.000000	400.000000	400.000000	396.000000	400.000000
mean	1.910000	1.162500	3.970000	2.565657	2.787500
std	1.611954	1.978859	4.215272	1.276162	1.822264
min	0.000000	0.000000	0.000000	1.000000	0.000000
25%	1.000000	0.000000	1.000000	2.000000	1.000000
50%	2.000000	0.000000	2.000000	2.000000	3.000000
75%	3.000000	1.000000	7.000000	4.000000	5.000000
max	8.000000	10.000000	18.000000	5.000000	5.000000

DATA.ISNULL.SUM

```
data.isnull().sum()
```

date	0
hair_loss	0
stay_up_late	0
pressure_level	0
coffee_consumed	0
brain_working_duration	0
school_assessment	305
stress_level	0
shampoo_brand	0
swimming	0
hair_washing	0
hair_grease	4
dandruff	289
libido	0
dtype: int64	

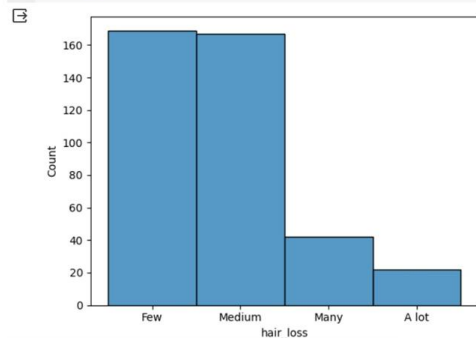
DATA DESCRIBE

```
data.describe()
```

	stay_up_late	coffee_consumed	brain_working_duration	hair_grease	libido
count	400.000000	400.000000	400.000000	396.000000	400.000000
mean	1.910000	1.162500	3.970000	2.565657	2.787500
std	1.611954	1.978859	4.215272	1.276162	1.822264
min	0.000000	0.000000	0.000000	1.000000	0.000000
25%	1.000000	0.000000	1.000000	2.000000	1.000000
50%	2.000000	0.000000	2.000000	2.000000	3.000000
75%	3.000000	1.000000	7.000000	4.000000	5.000000
max	8.000000	10.000000	18.000000	5.000000	5.000000

HISTPLOT

```
sns.histplot(x='hair_loss', data=data)
plt.show()
```



V DATA SET DESCRIPTION

To broaden a better pressure-primarily based hair fall detection and prevention system the use of K-Nearest Neighbours (KNN) and device gaining knowledge of strategies, you would need a dataset that captures different factors associated with pressure, hair fall, and other applicable capabilities. Here's a description of the dataset you may use:

Dataset Description:

Stress Levels:

Quantitative degree of stress degrees reported with the aid of individuals.

Collected through surveys, self-reporting apps, or wearable gadgets that degree physiological indicators of strain (e.g., heart price variability, pores and skin conductance).

Hair Fall Patterns:

Quantitative degree of hair fall, consisting of the range of strands shed in step with day or per week.

Collected through direct commentary or self-reporting with the aid of individuals.

Lifestyle Factors:

Sleep styles: Duration of sleep consistent with night, pleasant of sleep (e.g., measured via sleep monitoring devices).

Diet: Information approximately nutritional conduct, consumption of essential nutrients for hair health (e.g., protein, nutrients).

Exercise: Frequency and intensity of bodily activity.

Smoking and alcohol consumption: Quantitative measures of smoking and alcohol intake, if relevant.

Psychological Factors:

Anxiety and melancholy stages: Quantitative measures obtained thru standardized mental exams or self-reporting.

Coping mechanisms: Strategies individuals use to cope with strain (e.g., mindfulness practices, social help).

Demographic Information:

Age, gender, profession, and different demographic variables that may have an impact on pressure levels and hair fall styles.

Environmental Factors:

Pollution stages: Air and water great indices of the vicinity in which the character is living.

Climate: Temperature, humidity, and other weather-associated variables.

Medical History:

Any pre-existing scientific situations associated with hair fitness (e.g., alopecia, scalp problems).

Medications: Information approximately medications that individuals are taking, as a few medicines can purpose hair loss as a facet effect.

Outcome Variable:

Binary variable indicating the presence or absence of sizable hair fall, based totally on predefined standards (e.g., exceeding a positive threshold of hair strands shed in line with day).

Data Pre-processing:

Handling missing values: Address any missing values inside the dataset via imputation or deletion.

Feature scaling: Normalize or standardize numerical features to make certain they have a similar scale.

Encoding specific variables: Convert specific variables into numerical representations using strategies like one-hot encoding.

Feature choice: Identify and pick out relevant features that contribute maximum to the prediction of hair fall based on domain understanding and function significance evaluation.

With a nicely-organized dataset containing those features, you could train gadget gaining knowledge of fashions, together with KNN, to be expecting hair fall based on stress stages and other applicable factors. This can assist in growing effective prevention techniques tailored to people' wishes.

VI Design

DATA FLOW DIAGRAM:

1. DFD is likewise known as bubble table. It is a simple graphical formalism that may be used to symbolize the device in terms of the input facts to the machine, the diverse processing accomplished on those records, and the output statistics is carried out with the aid of that gadget.
2. A records drift diagram (DFD) is one of the most vital modelling tools. Used to model system additives. These additives are the device technique.
3. DFD indicates how statistics flows through the system and the way its miles converted through several transformations. It is a graphical method that represents the float of facts and the transformations that arise as facts moves from enter to output.

UML DIAGRAMS

This reputation is pushed and evolved through the Object Management Group.

The aim is that UML will subsequently come to be a commonplace language for designing item-oriented pc software models. A unified modelling language is a popular language for representing, visualizing, growing, and documenting artefacts of software program systems, as well as for modelling enterprise corporations and other non-software program systems.

UML represents a staple of nice-in-elegance engineering practices which have verified achievement in modelling huge and complicated structures.

UML is a crucial a part of element-oriented software program layout and software development method. UML by and large uses graphical notations to explain the presentation of software tasks.

GOALS:

The predominant layout targets of UML are:

1. Provide users with a demonstrative, usable visible modelling language with the intention to create and alternate significant models.
2. Provide extensibility and specialization mechanisms for extending centre standards.
3. Being unbiased of particular programming languages and development methods.
4. Provide a felony foundation for know-how the modelling language.

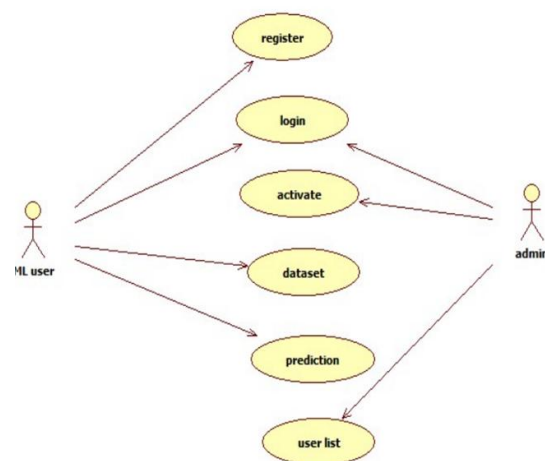
5. Encourage the growth of the OO tool market.

6. Support high-stage development concepts including relationships, frameworks, fashions and additives.

7. Integrate great practices.

USE CASE DIAGRAM:

A Unified Modelling Language (UML) use case diagram is a sort of behaviour diagram described and used in use case analysis. Its objective is to offer a graphical view of the functionalities furnished by way of the machine in phrases of actors, their goals (represented as use cases) and feasible dependencies among those use cases. The main goal of a use case diagram is to expose which device capabilities are used for which actor. The roles of the actors inside the machine may be represented.



VII MODEL ACCRACY TECHNIQUES

To enhance strain-primarily based hair fall detection and prevention the use of K-Nearest Neighbours (KNN) and device studying techniques, it is crucial to employ various accuracy enhancement strategies. Here's a rundown of strategies to enhance the accuracy of the gadget:

1. Feature Engineering:

Feature Selection: Identify the most applicable capabilities using strategies like correlation evaluation, feature significance rating, or area knowledge.

Feature Transformation: Transform functions the use of strategies like PCA (Principal Component Analysis) to lessen dimensionality while retaining essential records.

2. Data Augmentation:

Synthetic Data Generation: Generate synthetic facts points the use of strategies like SMOTE (Synthetic Minority Over-sampling Technique) to stability the dataset, specifically if the lessons are imbalanced.

3. Hyper parameter Tuning:

Grid Search or Random Search: Perform a systematic seek over a grid of hyper parameters or randomly sample hyper parameters to find the satisfactory combination that maximizes model performance.

4. Cross-Validation:

Use ok-fold go-validation to assess the version's overall performance across exceptional subsets of the statistics and make certain its generalization capability.

5. Ensemble Learning:

Voting Classifier: Combine multiple classifiers (which include KNN) and aggregate their predictions to improve ordinary performance.

Bagging (Bootstrap Aggregating): Train multiple instances of the classifier on exceptional subsets of the facts and combine their predictions to reduce variance.

Boosting: Sequentially educate vulnerable inexperienced persons, focusing on instances that have been misclassified with the aid of previous rookies, to enhance standard accuracy.

6. Model Evaluation Metrics:

Choose appropriate evaluation metrics depending on the trouble at hand. For binary category issues like hair fall detection, metrics which includes accuracy, precision, take into account, F1-score, and ROC-AUC score can offer insights into model overall performance.

7. Error Analysis:

Analyze the styles of mistakes made by the model (e.g., fake positives, fake negatives) to discover regions for improvement. This might also contain inspecting misclassified

times and understanding the reasons in the back of mispredictions.

8. Model Interpretability:

Use techniques inclusive of SHAP (Shapley Additive explanations) values or LIME (Local Interpretable Model-agnostic Explanations) to interpret the version's predictions and apprehend the contribution of every function to the very last selection.

9. Transfer Learning:

Transfer know-how from pre-trained fashions (if to be had) or leverage feature representations found out from related obligations to improve the overall performance of the hair fall detection version.

10. Regularization Techniques:

Apply regularization strategies like L1 or L2 regularization to prevent over fitting and improve the model's capacity to generalize to unseen statistics.

By using those accuracy enhancement strategies in mixture with KNN and other device gaining knowledge of strategies, you could increase a more desirable pressure-based hair fall detection and prevention device that offers dependable predictions and personalised tips for individuals.

K-Nearest Neighbours (KNN):

KNN is a supervised system gaining knowledge of algorithm used for each type and regression obligations.

It is one of the most simple but important algorithms in device mastering.

Key features of KNN:

Non-parametric: It doesn't anticipate any precise facts distribution.

Flexible: Handles both numerical and specific data.

Local shape-based totally: Makes predictions based on the similarity of facts factors inside the dataset.

Less touchy to outliers as compared to different algorithms.

How it works:

Given a new statistics factor, KNN unearths the K nearest buddies primarily based on a distance metric (normally Euclidean distance).

The elegance (for category) or cost (for regression) of the information point is determined through the general public vote or common of the K buddies.

Applications: Pattern popularity, records mining, intrusion detection, and extra.

Accuracy Techniques for KNN:

To enhance KNN accuracy, consider the subsequent techniques:

Feature Selection: Choose applicable capabilities and discard irrelevant ones. High-dimensional records can cause the "curse of dimensionality."

Distance Metrics: Experiment with different distance metrics (e.g., Euclidean,

Manhattan, Murkowski) to find the most appropriate one for your dataset.

Hyper parameter Tuning: Optimize hyper parameters just like the quantity of buddies (K) and distance weights.

Cross-Validation: Use techniques like okay-fold cross-validation to assess model performance.

Handling Imbalanced Data: Address class imbalance through oversampling, under sampling, or the usage of artificial statistics.

Weighted Voting: Assign one-of-a-kind weights to pals based totally on their proximity.

Ensemble Methods: Combine a couple of KNN fashions (e.g., bagging, boosting) for higher accuracy.

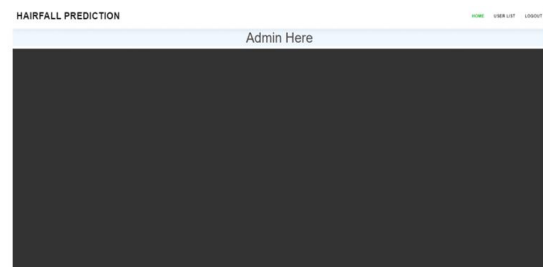
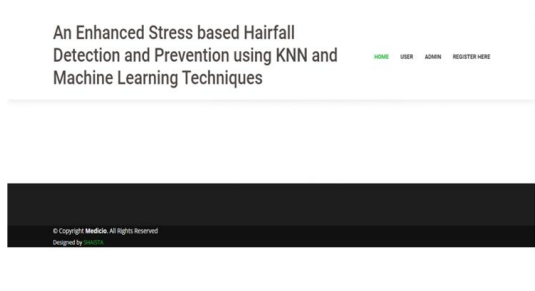
Distance Normalization: Normalize features to ensure identical importance.

Outlier Detection and Removal: Outliers can drastically have an effect on KNN; recall strong methods.

Distance-Weighted Voting: Weight buddies based totally on their distance from the query factor.

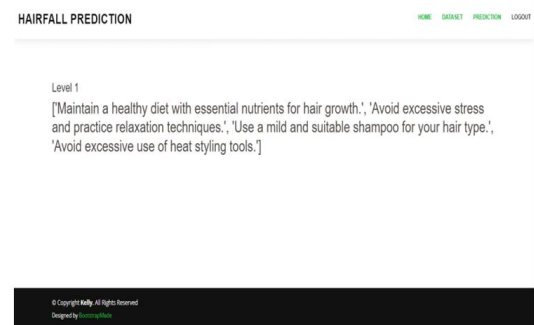
HOME PAGE:

USER REGISTER PAGE:



RESULTS:

Level1:



VIII CONCLUSION

An AI-based approach can allow for faster diagnosis. In this work, a machine gaining knowledge of set of rules changed into advanced to reliably predict 3 hair kinds. The proposed records set is likewise searched the usage of this approach. With the proposed method, the 3 maximum generic hair and scalp issues could gain from number one treatment choices and a higher expertise of the class of situations by means of physicians and sufferers. Therefore, the usages of hair photographs, a class gadget for regular hair and alopecia aerate become proposed.

REFERENCES

1. "Deep Learning based Detection of Hair Loss Levels from Face Pictures", 2019 Eighth International Conference on Image Processing Theory Tools and Applications, Halim Benhabiles, Karim Hammoudi, Ziheng Yang, Feryal Windal, Mahmoud Melkemi, Fadi Dornaika, and Ignacio Arganda-Carreras (IPTA)
2. Development of AI Hair Follicle Detection System and Associated Biomedical Goods, 2021 International Conference on Networking Systems of AI, Xin Zhang, Ruonan Zheng, Jinwen Lin, Yanru Zeng, and Yuming Zheng (INSAI)
3. "Survey-based Machine Learning approaches to diagnosis of hair fall disorder in Bangladeshi Community," 2022 13th International Conference on Computing Communication and Networking Technologies, Farhana Khatun, Moshfiqur Rahman Ajmain, Sharun Akter Khushbu, Nushrat Jahan Ria, and Sheak Rashed Haider Noori (ICCCNT)
4. A Deep Learning-Based Scalp Hair Inspection and Diagnostic System for Scalp Health, IEEE Access "Survey-based Machine Learning approaches to diagnosis of hair fall disorder in Bangladeshi Community," 2022 13th International Conference on Computing

Communication and Networking Technologies, Farhana Khatun, Moshfiqur Rahman Ajmain, Sharun Akter Khushbu, Nushrat Jahan Ria, and Sheak Rashed Haider Noori (ICCCNT)

5. Prediction of Alopecia Areata Using Machine Learning Approaches, 2022 IEEE International Conference on Data Science and Information System, S Aditya, Sanah Sidhu, and M Kanchana (ICDSIS)

6. An Intelligent Scalp Inspection and Diagnosis System for Caring for Hairy Scalp Health was developed by Jian-Ping Su, Liang-Bi Chen, Chia-Hao Hsu, Wei-Chien Wang, Cheng-Chin Kuo, Wan-Jung Chang, Wei-Wen Hu, and Da-Huei Lee and presented at the 2018 IEEE 7th Global Conference on Consumer Electronics (GCCE)

7. The authors are Hasanzadeh H, Nasrollahi S, Halavati N, Saberi M, and Firooz A. Male pattern hair loss treatment using 5% minoxidil topical foam: effectiveness, safety, and patient satisfaction. 2016; 25(3):41-44. Acta Dermatovenerol Alp Pannonica Adriat.

8. Investigation of the international epidemiology of androgenetic alopecia in young Caucasian men using images from the Internet, Avital Y, Morvay M,

Gaaland M, Kemny L. Indian Journal of Dermatology. 2015; 60(4):419.

9. Prasadu Peddi (2015) "A machine learning method intended to predict a student's academic achievement", ISSN: 2366-1313, Vol 1, issue 2, pp:23-37.