ML Models in HealthCare

HARIS KADU- UB 22009453
DANYALL AHMED -UB 21008895

Introduction

AI integration in healthcare has revolutionized disease diagnosis and treatment. Models like Naïve Bayes, Random Forest, and Decision Trees detect stroke, diabetes, and heart diseases with accuracy, saving lives. While these models enhance personalized care, they raise ethical concerns like data privacy and bias. This report examines methodologies, techniques, and ethical issues.

Naïve Bayes

Naïve Bayes predicts stroke using factors like blood pressure, smoking, and age. It handles categorical data well but struggles with correlated variables like smoking and hypertension. Hybrid models can improve its accuracy by addressing this limitation.

Decision Trees

Decision Trees are effective for diagnosing diabetes, classifying patients based on features like blood sugar, BMI, and family history. They can overfit noisy data, but pruning or limiting tree depth resolves this issue.

Random Forest

Random Forest improves heart disease predictions by combining decision trees for accuracy and reliability. It analyzes key variables like blood pressure, cholesterol, and ECG readings but may lack interpretability in scenarios needing clear judgments.

DS2 Image ML

A Convolutional Neural Network (CNN) is a type of deep learning model used to process images and similar data. It works by identifying patterns like edges and shapes using convolutional layers, simplifying the data with pooling layers, and making predictions with fully connected layers. CNNs are powerful for tasks like recognizing objects in images or diagnosing diseases from medical scans.

DS2 Sound ML

A Variational Autoencoder (VAE) is a type of neural network that learns to compress data into a simpler representation and then recreate it. Unlike regular autoencoders, VAEs encode data as a range of possible values, allowing them to generate new data by sampling from this range. They are often used for tasks like creating new images, reducing data dimensions, and exploring patterns in data.

Data Collection

Stroke – Native Bayes

The model predicts stroke based on age, hypertension, cholesterol, BMI, smoking, physical activity, nutrition, alcohol use, general and mental health, physical health, and family history of heart disease. These are important variables that contribute to the calculation of the risk of stroke.

Decision Trees – Diabetes

History of heart disease, hypertension, high cholesterol, BMI, physical activity, diet, any walking difficulty, age, education, and income are some of the features that this model uses to predict diabetes. Since machine learning models cannot read unprocessed text, categorical data needs to be transformed into a binary structure. Hypertension and body mass index are important markers because of their close relation to insulin resistance.

Random Forest - heart disease

Heart disease is influenced by factors such as blood pressure, cholesterol, BMI, diabetes, physical activity, diet, alcohol use, age, gender, and mental and physical health. High blood pressure damages arteries, increasing the risk of cardiac strain and heart failure. Elevated cholesterol accelerates heart disease progression, while a high BMI adds stress to the heart.

DS2 Image

A Convolutional Neural Network (CNN) detects pneumonia by analysing chest X-ray images. It identifies features like edges and patterns with convolutional layers, simplifies data through pooling layers, and uses fully connected layers to make predictions. The model determines the presence of pneumonia, offering an efficient and accurate diagnostic tool.

DS2 Sound

A Variational Autoencoder (VAE) supports pneumonia detection by compressing and reconstructing chest X-ray images. The encoder extracts key features into a compact form, while the decoder recreates the images. VAEs can generate synthetic X-rays for training, enhance noisy or low-quality images, and identify critical patterns.

Native Bayes

The dataset is loaded into a pandas DataFrame for manipulation, and key features were selected: Age, High Blood Pressure, High Cholesterol, BMI, Smoker, and Stroke as the target.

Splitting Data and Training Models

The data is split using train_test_split, reserving 80% for training and 20% for testing, ensuring consistent results with a defined random state. A Gaussian Naïve Bayes classifier, assuming features follow a normal distribution, is trained using X_train and y_train to identify correlations and predict outcomes effectively.

```
[[40793  5171]
 [ 2622  2150]]
              precision    recall  f1-score   support

           0       0.94      0.89      0.91     45964
           1       0.29      0.45      0.36      4772

    accuracy                           0.85     50736
   macro avg       0.62      0.67      0.63     50736
weighted avg       0.88      0.85      0.86     50736
```

Analysis

True negatives (40793): Correctly identified as No Stroke

False Positives (5171): Incorrectly identified as Stroke when No Stroke

False Negatives (2622): Incorrectly identified as No Stroke when Stroke

True Positives (2150): Correctly identified as Stroke

The Naïve Bayes model performs well in identifying non-stroke cases, achieving high precision (94%) and recall (89%) for the majority class. However, it struggles with stroke detection, showing low precision (29%) and recall (45%), resulting in numerous false positives and false negatives. The low F1 score of 0.36 for "Stroke" highlights the challenge of accurately identifying stroke cases.

Improve

Class Imbalance: For better representation while training, use techniques such as class weighting or oversample the minority class

Feature engineering in adding or improving features that could relate more strongly to incidents of stroke.

Decision Trees

Data Preparation

Matplotlib is used for visualization, NumPy for numerical operations, and pandas for data processing. Diabetes is the target variable (y), while Heart Disease or Attack, HighBP, and HighChol are predictors (X). Categorical variables like age, income, and education are converted to numerical values for machine learning compatibility. The dataset is split into 80% training and 20% testing using train_test_split with a fixed random state of 0 for reproducibility. A Decision Tree Classifier is trained to make predictions and evaluate performance on the test set.

Analysis

```
[[38646  3971]
 [ 5912  2207]]
              precision    recall  f1-score   support

           0       0.87      0.91      0.89     42617
           1       0.36      0.27      0.31      8119

    accuracy                           0.81     50736
   macro avg       0.61      0.59      0.60     50736
weighted avg       0.79      0.81      0.79     50736
```

True Negatives (38646): Correctly predicted as "No Diabetes."

False Positives (3971): Those were predicted as "Diabetes," whereas were "No Diabetes."

False Negatives (5912): The cases which are "Diabetes" in class but predicted as "No Diabetes."

True Positives (2207): Correctly predicted as those who have Diabetes.

<u>Summary</u>

The model achieves 81% accuracy but struggles to detect diabetes cases, with a low recall of 27%, indicating many undetected cases. Improving recall and precision would reduce false positives and enhance diabetes detection. Maximizing the F1 score should be a priority to improve overall performance. Recall can be improved through techniques like class balancing using under-sampling or over-sampling methods.

Random Forrest

<u>Data Preparation Heart Disease</u>

The dataset is pre-processed using pandas for data handling, matplotlib for visualization, and NumPy for computations. The target variable (y) is labeled as "1" for Heart Disease presence and "0" for absence. Key factors include HighBP, HighChol, and PhysActivity. Missing values are imputed using the mode for categorical features and the median for numerical ones. Features like age and BMI are scaled to improve model learning. The data is split into 80% training and 20% testing using train_test_split, and a Random Forest classifier is trained for predictions.

Analysis

```
[[44716  1248]
 [ 4111   661]]
              precision    recall  f1-score   support

           0       0.92      0.97      0.94     45964
           1       0.35      0.14      0.20      4772

    accuracy                           0.89     50736
   macro avg       0.63      0.56      0.57     50736
weighted avg       0.86      0.89      0.87     50736
```

True Negatives (44716): 44,716 patients were accurately predicted by the model as having No Heart disease.

False Positives (1248): 1,248 cases were misclassified by the model as Heart Disease.

False Negatives (4,111): The model failed to identify 4,111 instances of true Heart Disease.

Accuracy for the Heart Disease Prediction model is 89.4%. This performs very well in Class 0, which indicates No Heart Disease, while performing terribly in the Class 1 Heart Disease.

- Class 0 precision = 92%.

- Class 1 precision = 35%

- Class 0 recall = 97% where Class 1 recall = 14%

F1 Score : Class 0 : 0.94; Class 1 : 0.20.

Class imbalance: The solution will be the addressing of class imbalance using class weighting or resampling strategies in order to improve the detection of heart disease.

While the model is relatively good at predicting a no heart illness outcome, it does need to improve significantly to be able to predict the presence of heart disease correctly.

Image ML Model

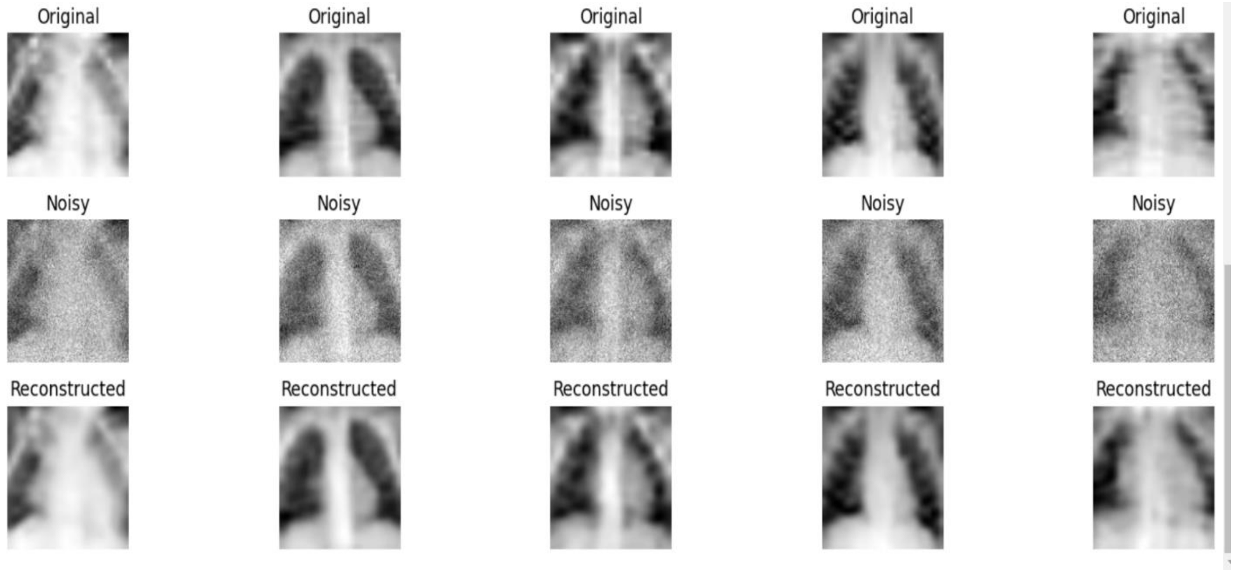
```
148/148 ━━━━━━━━━━━━━━━━━━━━ 158s 1s/step - accuracy: 0.8756 - loss: 0.7093 - val_accuracy: 0.6250 - val_loss:
Epoch 2/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 152s 1s/step - accuracy: 0.9388 - loss: 0.2104 - val_accuracy: 0.6250 - val_loss: 4.7346
Epoch 3/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 151s 1s/step - accuracy: 0.9518 - loss: 0.1335 - val_accuracy: 0.6987 - val_loss: 2.0340
Epoch 4/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 150s 1s/step - accuracy: 0.9533 - loss: 0.1384 - val_accuracy: 0.7949 - val_loss: 1.1212
Epoch 5/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 150s 1s/step - accuracy: 0.9534 - loss: 0.1431 - val_accuracy: 0.8221 - val_loss: 0.6327
Epoch 6/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 156s 1s/step - accuracy: 0.9590 - loss: 0.1356 - val_accuracy: 0.8365 - val_loss: 0.9857
Epoch 7/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 152s 1s/step - accuracy: 0.9683 - loss: 0.0949 - val_accuracy: 0.8141 - val_loss: 1.3729
Epoch 8/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 150s 1s/step - accuracy: 0.9641 - loss: 0.1034 - val_accuracy: 0.8237 - val_loss: 1.2471
Epoch 9/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 152s 1s/step - accuracy: 0.9674 - loss: 0.0802 - val_accuracy: 0.8702 - val_loss: 0.9522
Epoch 10/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 157s 1s/step - accuracy: 0.9725 - loss: 0.0691 - val_accuracy: 0.8702 - val_loss: 0.9420
```

The CNN model performed well with 97% training accuracy and 87% validation accuracy, though it shows some overfitting. Improving it with more diverse data and techniques like data augmentation can help. This tool can speed up pneumonia detection, helping doctors diagnose faster. However, it's important to protect patient data, avoid bias, and use the AI as a support tool, not a replacement for doctors, to ensure safe and fair use.

Sound ML Model

```
Training the classification model...
Epoch 1/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 154s 1s/step - accuracy: 0.8723 - loss: 1.0433 - val_accuracy: 0.6250 - val_loss: 19.6354
Epoch 2/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 202s 1s/step - accuracy: 0.9341 - loss: 0.1853 - val_accuracy: 0.6250 - val_loss: 14.8949
Epoch 3/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 201s 1s/step - accuracy: 0.9462 - loss: 0.1399 - val_accuracy: 0.6266 - val_loss: 4.8795
Epoch 4/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 147s 992ms/step - accuracy: 0.9541 - loss: 0.1197 - val_accuracy: 0.7500 - val_loss: 1.3953
Epoch 5/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 148s 1s/step - accuracy: 0.9579 - loss: 0.1055 - val_accuracy: 0.7885 - val_loss: 0.8481
Epoch 6/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 203s 1s/step - accuracy: 0.9669 - loss: 0.1078 - val_accuracy: 0.5721 - val_loss: 5.3201
Epoch 7/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 202s 1s/step - accuracy: 0.9604 - loss: 0.1021 - val_accuracy: 0.8494 - val_loss: 1.0040
Epoch 8/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 202s 1s/step - accuracy: 0.9662 - loss: 0.0891 - val_accuracy: 0.8397 - val_loss: 0.9686
Epoch 9/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 200s 1s/step - accuracy: 0.9629 - loss: 0.1057 - val_accuracy: 0.7772 - val_loss: 1.3639
Epoch 10/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 202s 1s/step - accuracy: 0.9657 - loss: 0.0856 - val_accuracy: 0.8285 - val_loss: 0.7876
Epoch 11/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 149s 1s/step - accuracy: 0.9659 - loss: 0.0940 - val_accuracy: 0.8590 - val_loss: 0.6030
Epoch 12/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 150s 1s/step - accuracy: 0.9681 - loss: 0.0842 - val_accuracy: 0.8494 - val_loss: 0.6725
Epoch 13/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 202s 1s/step - accuracy: 0.9727 - loss: 0.0717 - val_accuracy: 0.8766 - val_loss: 0.6679
Epoch 14/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 200s 996ms/step - accuracy: 0.9737 - loss: 0.0651 - val_accuracy: 0.8510 - val_loss: 0.8418
Epoch 15/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 203s 1s/step - accuracy: 0.9775 - loss: 0.0561 - val_accuracy: 0.8125 - val_loss: 1.0450
Epoch 16/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 201s 995ms/step - accuracy: 0.9735 - loss: 0.0788 - val_accuracy: 0.4519 - val_loss: 7.2471
Epoch 17/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 202s 1s/step - accuracy: 0.9759 - loss: 0.0642 - val_accuracy: 0.8862 - val_loss: 0.9687
Epoch 18/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 148s 1s/step - accuracy: 0.9780 - loss: 0.0583 - val_accuracy: 0.3750 - val_loss: 63.3907
Epoch 19/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 147s 994ms/step - accuracy: 0.9761 - loss: 0.0889 - val_accuracy: 0.8205 - val_loss: 1.3655
Epoch 20/20
```

```
Epoch 1/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 391s 3s/step - loss: 0.0123 - val_loss: 6.1382e-04
Epoch 2/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 449s 3s/step - loss: 5.0698e-04 - val_loss: 5.6808e-04
Epoch 3/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 439s 3s/step - loss: 4.3032e-04 - val_loss: 3.1374e-04
Epoch 4/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 435s 3s/step - loss: 3.4433e-04 - val_loss: 4.4288e-04
Epoch 5/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 442s 3s/step - loss: 3.4425e-04 - val_loss: 2.9631e-04
Epoch 6/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 443s 3s/step - loss: 3.0991e-04 - val_loss: 2.7644e-04
Epoch 7/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 441s 3s/step - loss: 2.8156e-04 - val_loss: 3.2800e-04
Epoch 8/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 444s 3s/step - loss: 2.7964e-04 - val_loss: 2.5914e-04
Epoch 9/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 441s 3s/step - loss: 2.7305e-04 - val_loss: 6.3021e-04
Epoch 10/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 389s 3s/step - loss: 7.0802e-04 - val_loss: 2.9958e-04
Epoch 11/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 440s 3s/step - loss: 2.4872e-04 - val_loss: 2.6401e-04
Epoch 12/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 443s 3s/step - loss: 2.3744e-04 - val_loss: 2.4066e-04
Epoch 13/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 441s 3s/step - loss: 2.4399e-04 - val_loss: 2.8486e-04
Epoch 14/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 442s 3s/step - loss: 2.4672e-04 - val_loss: 3.0052e-04
Epoch 15/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 443s 3s/step - loss: 2.9090e-04 - val_loss: 2.3675e-04
Epoch 16/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 388s 3s/step - loss: 2.3009e-04 - val_loss: 2.5597e-04
Epoch 17/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 441s 3s/step - loss: 2.3521e-04 - val_loss: 2.3141e-04
Epoch 18/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 386s 3s/step - loss: 2.2346e-04 - val_loss: 2.2212e-04
Epoch 19/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 442s 3s/step - loss: 2.2126e-04 - val_loss: 3.6444e-04
Epoch 20/20
148/148 ━━━━━━━━━━━━━━━━━━━━ 443s 3s/step - loss: 3.3795e-04 - val_loss: 2.1697e-04
Visualizing denoising results...
```



The CNN classification model achieved 97.6% training accuracy but showed overfitting, with validation accuracy around 88.6% and spikes in validation loss. The generative model effectively denoised noisy X-rays, producing clean reconstructions with a low validation loss of 0.000216. These tools accelerate pneumonia detection by automating diagnosis and enhancing image quality, but ethical concerns like data privacy, bias, and AI reliance must be addressed for fair and safe use in healthcare.

Ethical Issues

Ethical concerns include maintaining user privacy by adhering to data protection regulations and treating data as a security priority. Developers must minimize unexpected outcomes through rigorous testing and ensure models are inclusive, addressing the needs of underrepresented populations. Upholding these principles enhances productivity, societal welfare, and adherence to ethical machine learning standards.

Conclusion

AI models like Naïve Bayes, Decision Trees, Random Forest, CNNs, and VAEs demonstrate great potential in improving disease detection and patient monitoring. However, addressing challenges like overfitting, class imbalance, and ethical concerns, including data privacy and fairness, is essential. With proper improvements and safeguards, these models can revolutionize healthcare by making diagnosis faster, more accurate, and widely accessible.

Bibliography

- Brown, S. (2021). *Machine learning, explained*. [online] MIT Sloan. Available at: https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained.
- Kanade, V. (2022). *What Is Machine Learning? Definition, Types, Applications, and Trends for 2022*. [online] Spiceworks. Available at: https://www.spiceworks.com/tech/artificial-intelligence/articles/what-is-ml/.
- IBM (2021). *Naive Bayes*. [online] Ibm.com. Available at: https://www.ibm.com/think/topics/naive-bayes.
- Castillo, D. (2021). *Decision trees in machine learning explained*. [online] Seldon. Available at: https://www.seldon.io/decision-trees-in-machine-learning.
- Institute of Data (2024). *Demystifying Random Forests: A Comprehensive Guide | Institute of Data*. [online] Institute of Data. Available at: https://www.institutedata.com/blog/random-forests-a-guide/.
- Jain, A. (2024). *Everything about Random Forest - Abhishek Jain - Medium*. [online] Medium. Available at: https://medium.com/@abhishekjainindore24/everything-about-random-forest-90c106d63989.
- Analytics Vidhya. (2020). *CNN Image Classification | Image Classification Using CNN*. [online] Available at: https://www.analyticsvidhya.com/blog/2020/02/learn-image-classification-cnn-convolutional-neural-networks-3-datasets/.
- Sutaria, N. (2022). *Bias and Ethical Concerns in Machine Learning*. [online] ISACA. Available at: https://www.isaca.org/resources/isaca-journal/issues/2022/volume-4/bias-and-ethical-concerns-in-machine-learning.
- ecanorea (2024). *Ethics and Machine Learning: Present and Future Challenges - Plain Concepts*. [online] Plain Concepts. Available at: https://www.plainconcepts.com/ethics-machine-learning-challenges/.