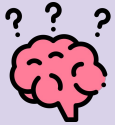
Project 4, Group 9

Alzheimer's Disease Prediction

Using Data Science to Make Early Detection More Accessible

Presented by: Luisa Murillo, Monica Mitry, Gorgina Kareem, & Adriana Kuhl

Inspiration and Purpose :



Importance of Early Diagnosis in Alzheimer's Disease:Emphasize that Alzheimer's disease is a significant global health issue, and early detection is crucial for improving patient care and treatment outcomes.

- 6 million Americans live with Alzheimer’s
- Diagnosis is expensive and requires clinical access

Use of Machine Learning in Healthcare: Discuss how machine learning techniques can be applied to predict the likelihood of Alzheimer's disease, helping improve diagnosis accuracy and treatment strategies.

Inspiration from Data-Driven Healthcare Solutions: Highlight the growing role of data-driven approaches in healthcare, particularly for diagnosing and managing complex diseases like Alzheimer's, and how this project aims to explore predictive analytics for earlier intervention.

Dataset and Cleaning:

For our dataset we used the “Alzheimer’s Disease Prediction” found on Kaggle by Adham Tarak and Anna Balatska

Dropped columns.

Refined dataset for analytical clarity.

Data mapping for certain columns.

Research Questions:

- Age and Alzheimer’s Onset: Is there an average age that Alzheimer's symptoms begin to appear?

- Demographics and Risk: Is there a correlation between ethnicity and a higher risk of Alzheimer’s?

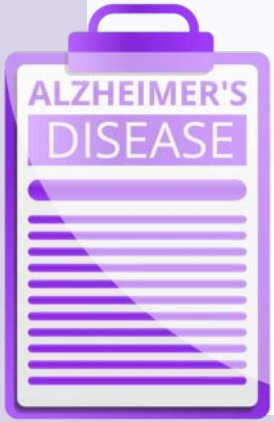
- Lifestyle & Health Correlations: Does a patient's lifestyle play a role in the risk of Alzheimer’s?

Machine Learning:

- **Preprocessing steps**: (e.g., scaling, and label encoding) were applied to prepare the data for modeling.
    - For label encoding we changed the columns of “Gender” and “Ethnicity” from strings to integers to have the model process the data more smoothly.
- **Four ML models tested**: Linear Regression, Decision Trees, Random Forest, and XGBoost.
    - Models were **evaluated based on AUC, F1-score, and accuracy** for Alzheimer’s diagnosis.
    - XGBoost and Random Forest showed the strongest potential for early detection.
        - Random Forest: AUC: 0.82, F1-score: 0.58, Accuracy: 0.72
        - XGBoost: AUC:0.80, F1-score: 0.57, Accuracy: 0.70

- **A Random Forest** model was chosen to train our model as it outperformed slightly in all categories.

Live Demo:


ALZHEIMER'S
DISEASE

Navbar
Home
About Us
Welcome to Alzheimers Predictor App
About This Project
This app is designed to be an interactive website to learn more about the causes and probability of developing Alzheimers
Alzheimer's Probability Predictor
Enter user data.

Data limitations & Bias:

- **Incomplete or missing data** may reduce the accuracy and reliability of the predictive model.
- **Demographic imbalance** can introduce bias as certain ethnicities, age groups, or income levels may be underrepresented
  .
- **Regional bias** may occur if data comes mostly from certain locations or healthcare systems with better access to diagnostics.
- **Diagnosis inconsistencies** can arise if medical evaluations or data collection methods vary across sources.
- **Limited generalizability** — the model might not perform equally well on populations outside the dataset.
- **Self-reported data** (if included) can be subjective and prone to inaccuracies.

Conclusion and Future Work:

ML model showed strong potential in predicting Alzheimer's diagnosis from patient data. The biggest correlation linked Alzheimer's for the model being the functional assessment outcome.

Visualizations revealed key patterns between age, gender, ethnicity, and family history. From Tableau we can visualize that Alzheimer's ranges from ages 60 and above.

There is a correlation between ethnicity and a higher risk of Alzheimer's. Given the data in our pie chart we find Alzheimer's to impact every ethnicity between 31%-41%, with Caucasian ethnicity having the higher percentage at 41%.

Lifestyle & Health Correlations: It was found the lifestyle choices have less impact in developing Alzheimer's later in life.

Future research can build on this approach to improve care and quality of life for Alzheimer's patients.

Questions?