

# Enhancing FEMA Disaster Response through Machine Learning by Predicting Rental Assistance Eligibility

## INTRODUCTION

Disasters like hurricanes and storms frequently leave survivors in urgent need of housing assistance. However, the current application process for FEMA's rental assistance program often takes 2 to 3 weeks to complete, which may lack a data-driven framework. To address this challenge, this research aims to create data-driven models using supervised machine learning algorithms . These models will predict rental assistance eligibility post-disaster, enabling FEMA decision-makers to expedite application procedures and provide applicants with quicker determinations. By leveraging a dataset of 6.3 million application records from the National Emergency Management Information System (NEMS), the research aims to develop classification models capable of accurately identifying eligible recipients.

Classification models - Logistic Regression, Decision Tree, KNN.



## METHODOLOGY

### Data Collection

The data is imported from FEMA open source(OpenFEMA) which has 6.3 million rows and 42 columns

### Data preprocessing

Handling Missing Values, Encoding Categorical Variables, Feature scaling, Datacleaning

### Feature engineering

Selecting relevant features by research, correlation analysis, feature importance ranking

### Model selection & Training

Choosing an appropriate algorithm for the classification task. Training the chosen model on labeled training data.

### Evaluation

Assessing the performance of the trained model on unseen data using appropriate metrics (e.g., accuracy, precision, recall, F1-score)

## ANALYSIS



The bar chart illustrates FEMA's expenditure of \$4,250.49 million on assistance programs, with repair receiving the highest allocation, indicating a focus on structural rehabilitation. However, a significant portion of the funds is also allocated to rental assistance, highlighting the agency's commitment to providing temporary housing solutions.

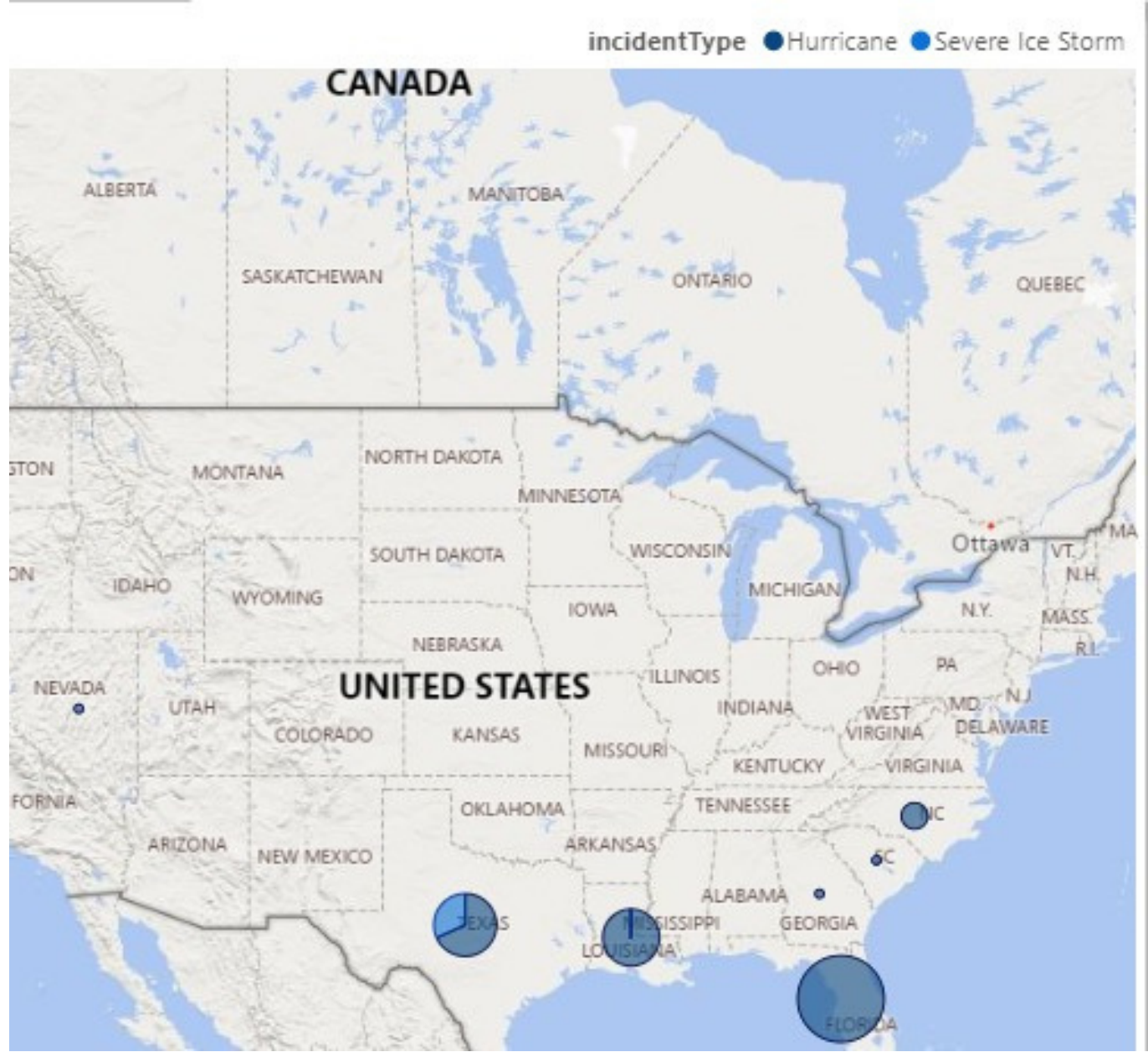
The scatter plots depict the joint distribution of three variables: household composition, gross income, and residence type. The majority of data points are clustered near the origin, suggesting limited correlation between these variables. However, outliers are observed, particularly in the gross income distribution, indicating a long-tail distribution. Additionally, the vertical line patterns seen when plotting residence type against the other variables suggest that residence type may be categorical in nature.

Based on our research and feature selection techniques below are independent variables to predict rental Assistance Eligible

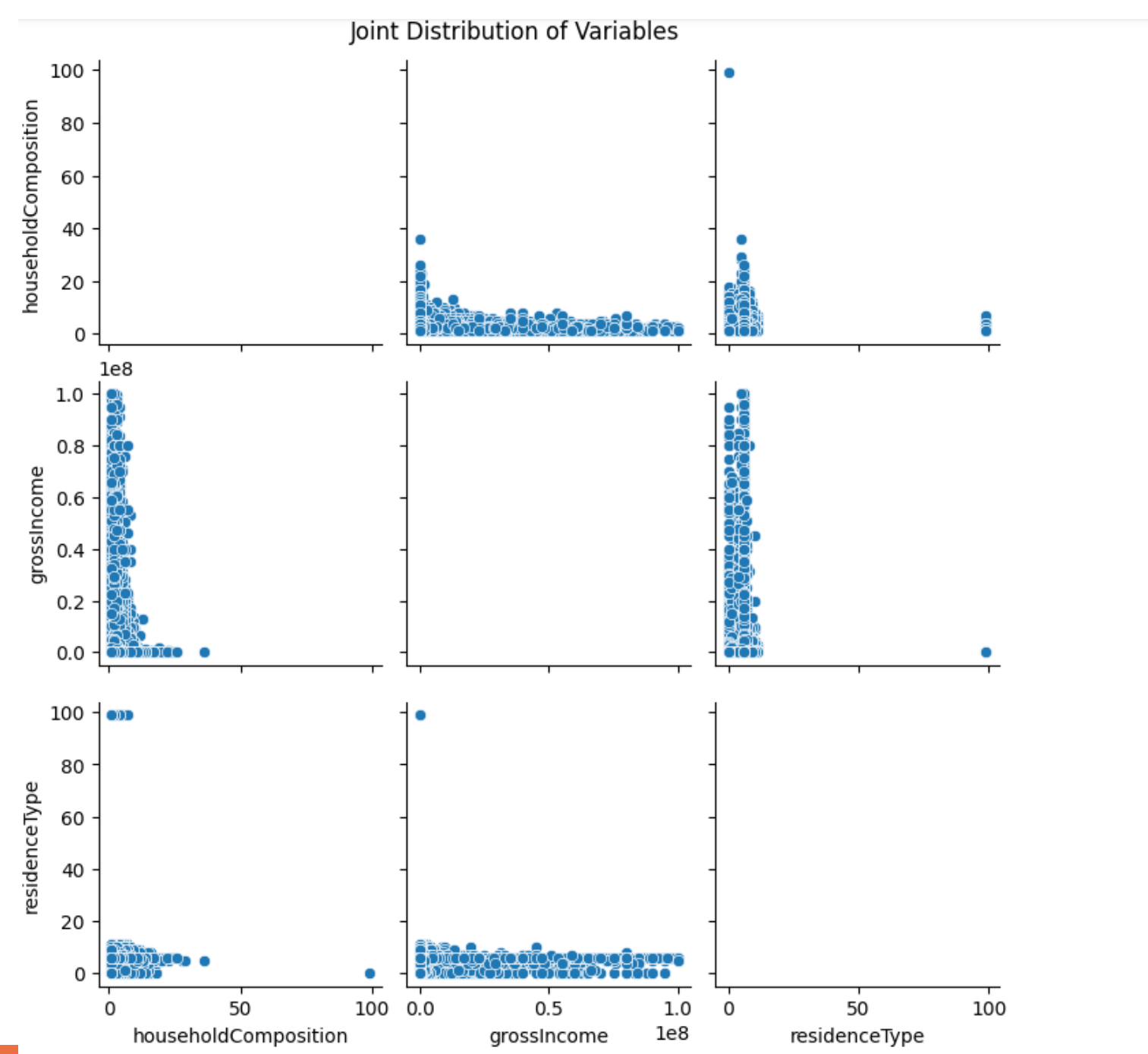
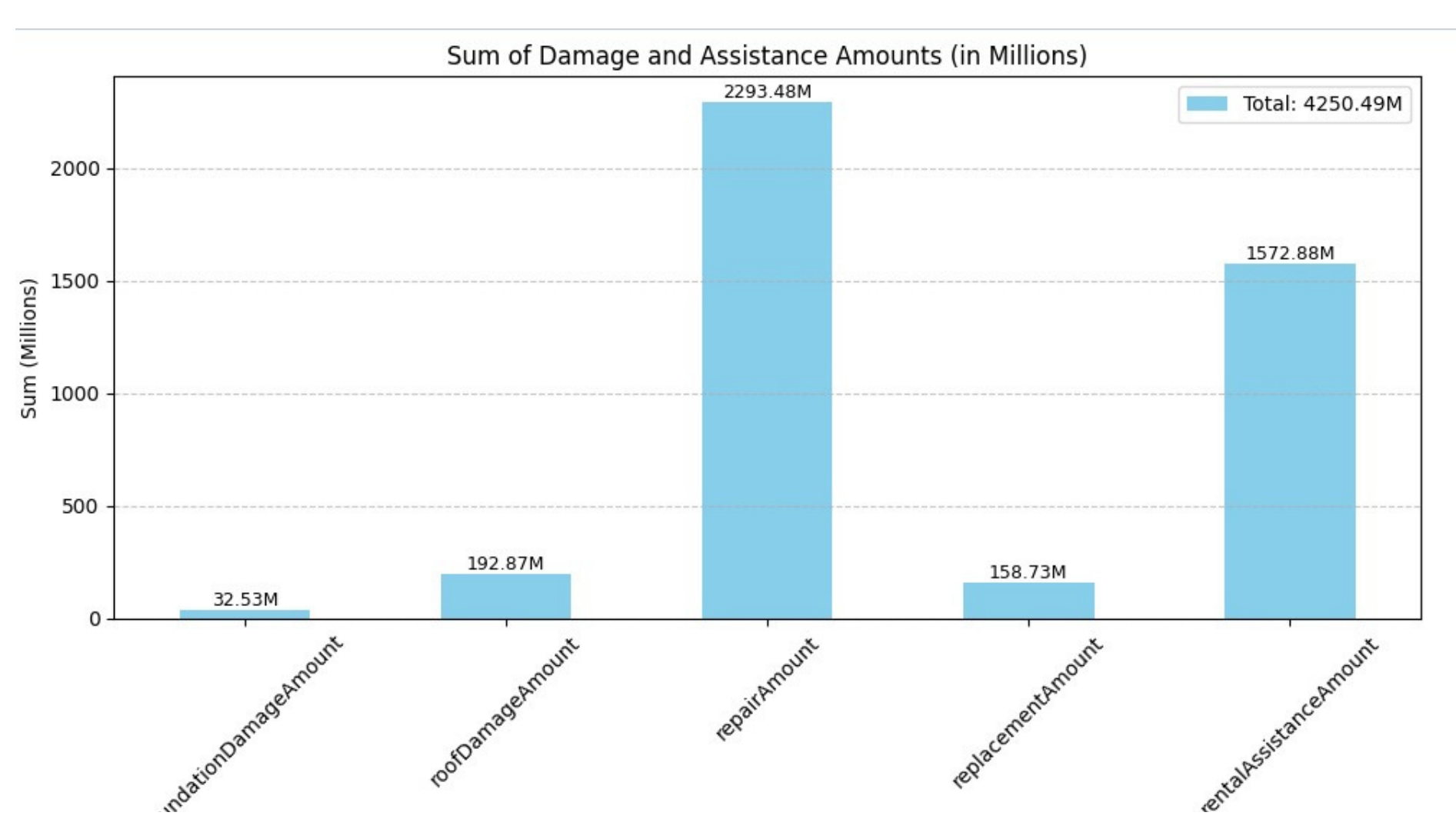
Independent variables – household Composition, gross Income, special Needs, own Rent, residence Type, homeowners Insurance, flood Insurance, inspected column, habitability Repairs Required, destroyed, water Level, flood Damage, foundation Damage, roof Damage.

Dependent variable - Rental Assistance Eligible

In terms of predictive modeling performance, logistic regression achieved an accuracy of 65.97%, with a sensitivity of 2.98% and a precision of 70.73%. The decision tree model demonstrated a slightly higher accuracy of 68.83%, but lower sensitivity and precision at 1.30% and 43.03%, respectively. On the other hand, the K nearest neighbor (KNN) model outperformed both with an accuracy of 70%, sensitivity of 16.80%, and precision of 45.61%.



| damagedStateAbbreviation | Hurricane | Severe Ice Storm |
|--------------------------|-----------|------------------|
| FL                       | 452332    |                  |
| TX                       | 147605    |                  |
| PR                       | 185279    |                  |
| LA                       | 171775    |                  |
| NC                       | 22946     |                  |
| SC                       | 3         |                  |
| GA                       | 2         |                  |
| NV                       | 1         |                  |



| Table 5. Accuracy values for each model |                    |                       |                           |
|---|--------------------|-----------------------|---------------------------|
| Model                                   | Accuracy (percent) | Sensitivity (percent) | Precision score (percent) |
| Logistic regression                     | 65.97              | 2.98                  | 70.73                     |
| Decision tree                           | 68.83              | 1.30                  | 43.03                     |
| KNN                                     | 70                 | 16.80                 | 45.61                     |

KNN: K nearest neighbor.

**Conclusion - The results of the research highlight the potential of data-driven strategies to improve FEMA's rental assistance program, especially in situations involving disaster relief. This research demonstrated methods to streamline the eligibility decision process by utilizing massive datasets and machine learning algorithms. This translates to quicker assistance for disaster survivors facing urgent housing needs. Furthermore, the potential development of a self-assessment tool based on these models could empower individuals to proactively assess their eligibility, facilitating faster disaster recovery efforts.**

### References :

Federal Emergency Management Agency (FEMA), OpenFEMA Dataset: Individual Assistance Housing Registrants - Large Disasters -v2. Retrieved from <https://www.fema.gov/openfema-data-page/individual-assistance-housing-registrants-large-disasters-v1> on March 6, 2024, 12:34 PM EST. This product uses the FEMA OpenFEMA API, but is not endorsed by FEMA. The Federal Government or FEMA cannot vouch for the data or analyses derived from these data after the data have been retrieved from the Agency's website(s).  
FEMA: Transitional shelter assistance fact sheet. 2017.- [https://www.researchgate.net/profile/Emad-Elwakil-2/publication/356960380\\_Machine\\_learning-based\\_FEMA\\_Transitional\\_Shelter\\_Assistance\\_TSA\\_eligibility\\_prediction\\_models/links/6266c661bca601538b679476/Machine-learning-based-FEMA-Transitional-Shelter-Assistance-TSA-eligibility-prediction-models.pdf](https://www.researchgate.net/profile/Emad-Elwakil-2/publication/356960380_Machine_learning-based_FEMA_Transitional_Shelter_Assistance_TSA_eligibility_prediction_models/links/6266c661bca601538b679476/Machine-learning-based-FEMA-Transitional-Shelter-Assistance-TSA-eligibility-prediction-models.pdf)

### AUTHORS

**STUDENTS- ROHITH SANKU, DURGA PRASANNA**  
**PROFESSOR- GULHAN BIZEL**