



MediMatch

Course Name: Machine Learning (UML501)

Subgroup: 3CO12

Semester: 2425ODDSEM

Project Name: MediMatch

Submitted To: Dr. Jyoti Maggu

Submitted By: Pranjal Kishor (102203495)

Gopal Agarwal (102203554)

THE PROBLEM



+ Limited Medicine Availability

Patients often face interruptions in treatment when prescribed medications are out of stock or unavailable in their region.

+ Challenges in identifying safe alternatives

Healthcare providers and patients struggle to find appropriate substitute medications, especially for conditions requiring specific formulations or for those with allergies.

+ Delays in Treatment

Delays in treatment due to lack of suitable prescription alternatives, impacting patient care and healthcare efficiency.



OUR INNOVATION: MediMatch

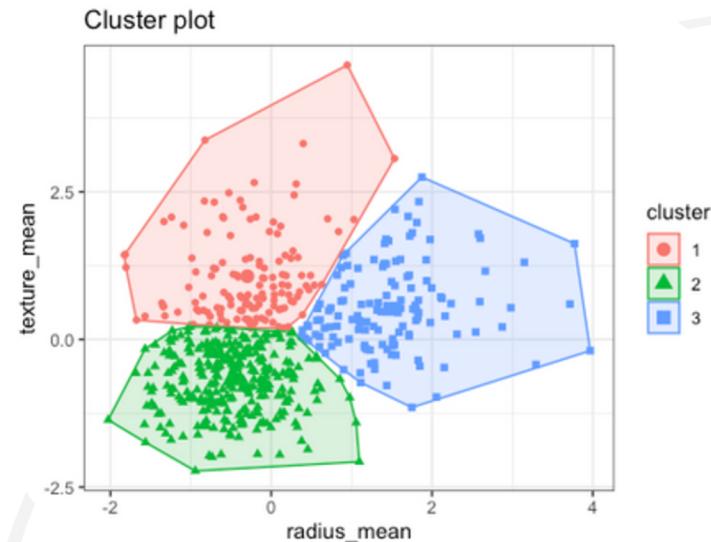


Real-Time Ball Tracking and Prediction.

MediMatch uses machine learning to identify effective, safe alternatives for allergies, aiding both patients and healthcare providers efficiently.

AR-Based 3D Models to Visualize Plays.

MediMatch minimizes treatment delays by providing alternative options, improving healthcare efficiency and supporting timely, effective patient care.



Provides alternate medication using advance ML algorithms



TECHNICAL WORKFLOW

MediMatch

• Data Cleaning



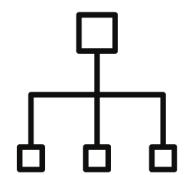
- Dropped unnecessary columns to streamline the dataset, focusing only on essential attributes.
- Handled missing values, ensuring the dataset is complete and ready for analysis.

• Feature Engineering



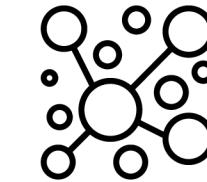
- Applied one-hot encoding on categorical columns to transform them into numerical form for compatibility with clustering algorithms.
- Standardized numerical features to ensure consistency in scale across attributes, which helps improve the accuracy of clustering

• Hierarchical Clustering



- Used hierarchical clustering with Ward's method, which minimizes variance within clusters, grouping similar medicines based on all relevant attributes.
- Generated a dendrogram to visually confirm the clustering structure .

• Cluster Assignment



- Assigned each medicine to a cluster based on hierarchical clustering results.
- Saved key model components (linkage matrix, cluster assignments) using Python's pickle module, making it easy to load the model in future sessions for real-time queries without re-computing clusters.

• Similarity Search and Recommendation



- Implemented a KNN-based search within each cluster to identify the most similar alternative for any given medicine. This allows for efficient and precise alternative recommendations.

Research & References

- Dataset Link

<https://www.kaggle.com/datasets/singhnavjot2062001/11000-medicine-details>

- Research paper used for references

-

F. Murtagh and P. Contreras, "Algorithms for hierarchical clustering: an overview," WIREs Data Mining and Knowledge Discovery, vol. 2, no. 1, pp. 86-97, 2012.

J. H. Ward, "Hierarchical grouping to optimize an objective function," Journal of the American Statistical Association, vol. 58, no. 301, pp. 236-244, 1963