

# Understanding How Doctors Judge

Analyzing Patient Characteristics to Predict Decision-Making

Noan Gutman, Rotem Eidlitz and Yuval Vinokur | Sarel Cohen and Eyal zinger

## BACKGROUND

The project involves a dataset of patient characteristics, with the ultimate goal of determining whether a patient has psychosis or neurosis. The scale ranges from 0 (psychosis) to 11 (neurosis). In addition to patient data, there are responses from 29 different doctors. source : Nonlinear Models of Clinical Judgment: Meehl's Data Revisited" by Yoav Ganzach from Tel Aviv University.



## MAIN GOALS

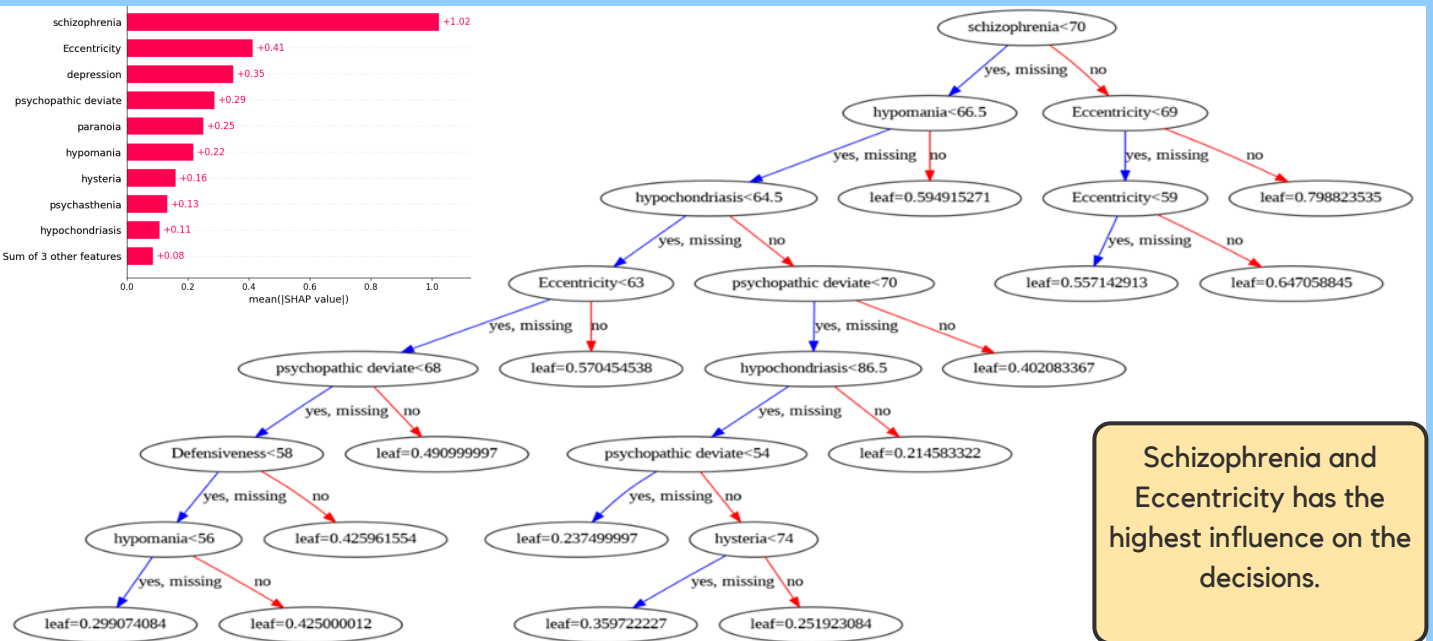
- Understand how doctors make their decision
  - Predict the decision of a specific doctor.
  - Predict the actual decision (psychosis or neurosis).

## ML MODELS

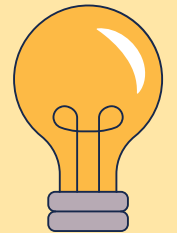
After training multiple models,(logistic, linear, forest) we have determined that XGBoost outperforms the others in predicting these decisions accurately.

ML model	forest	logistic	linear	XGBoost
R2 score	0.74619305	0.68071031	0.72479292	0.78350176

## ANALYZING XGBOOST MODEL



## EXPLORING CORRELATION BETWEEN JUDGES



we investigated the correlation between different judges in our dataset. As we aim to understand how doctors make their decisions, analysing the level of agreement or disagreement among them can provide valuable insights into the decision-making proces

Cluster 0 judges: 7 8 9 10 11 18 19 20 21 23 24 26  
 Cluster 1 judges: 0 2 4 5 12 15 27 28  
 Cluster 2 judges: 1 3 6 14 16 17 22  
 Outlier: 13

DBSCAN was employed to attain these clusters.

## UNIVERSAL JUDGE

We trained the universal judge by averaging decisions from all judges. Employing XGBoost, linear regression, and random forest techniques alongside the Universal judge, we predicted outcomes for each of the 29 judges. Subsequently, we gauged the judges' correlation with these predictions using the R2 score.

Conclusion:  
 Our application of the Universal judge model to predict judicial decisions yielded insightful results. Judges grouped within DBSCAN clusters demonstrated closely aligned R2 scores, reaffirming the model's ability to capture shared tendencies. Conversely, Judge 13, identified as an outlier within the DBSCAN clusters, exhibited the lowest score of 0.1. This study highlights the model's effectiveness in recognizing consistent decision patterns among clustered judges, along with its capability to identify exceptional cases such as Judge 13.

## CONCLUSION

Predicting the decisions of specific doctors from our dataset is challenging. It appears that each doctor follows an unpredictable decision-making process unique to them. Our investigation did not reveal any correlations among different doctors, leading us to conclude that no such connections exist.

