

## INTISARI

# PREDIKSI *ADMISSION SCORE* GAOKAO PADA TINGKAT PROGRAM STUDI DENGAN KOMBINASI *SUPERVISED LEARNING* DAN *K-MEANS CLUSTERING*

Oleh

Michael Rafel Hosea Siregar

22/508531/PPA/06439

Gaokao merupakan sistem seleksi perguruan tinggi berbasis tes di Tiongkok yang juga merupakan salah satu ujian tersulit di dunia. Pengembangan model prediksi *admission score* Gaokao pada tingkat program studi perlu dilakukan, tetapi penelitian yang ada saat ini baru melakukan prediksi pada tingkat universitas. Oleh karena itu, penelitian ini berfokus pada prediksi *admission score* Gaokao pada tingkat program studi dengan mempertimbangkan tingkat kompetisi antar program studi sekaligus menentukan model kompetisi yang paling tepat menggunakan *elbow method*.

Dalam penelitian ini, prediksi dilakukan dalam dua tahap, yaitu *preliminary prediction* dan *fine prediction*. *Preliminary prediction* dilakukan menggunakan beberapa metode, seperti *Gradient Boosted Decision Tree*, *Linear Regression*, *Multi Layer Perceptron*, dan LSTM. Setelah itu, setiap program studi di-*cluster* menggunakan K-Means Clustering yang didasarkan pada *admission score* pada tahun 2019 dan hasil prediksi *admission score* tahun 2020 unuk memperoleh *fine prediction*, yaitu rerata *admission score* program studi tersebut pada tahun 2020 berdasarkan hasil *preliminary prediction* dengan *admission score* universitas pada posisi yang sama pada tahun sebelumnya.

Metode terbaik dalam penelitian ini adalah *Gradient Boosted Decision Tree* yang memberikan *Mean Absolute Error* sebesar 5,1176, *Mean Square Error* sebesar 58,0890, dan *precision* sebesar 47,059% jika digunakan *threshold* sama dengan 3. Adapun metode *Long Short Term Memory* menghasilkan *Mean Square Error* sebesar 269,2584, *Mean Absolute Error* sebesar 14,4024, dan *precision* sebesar 9,756% jika digunakan *threshold* sama dengan 3.

**Kata Kunci:** *Decision Tree*, *Linear Regression*, *Multi Layer Perceptron*, *Long Short Term Memory*, Prediksi Skor Gaokao

## ABSTRACT

### ***GAOKAO ADMISSION SCORE PREDICTION IN MAJOR LEVEL WITH A COMBINATION OF SUPERVISED LEARNING AND K-MEANS CLUSTERING***

By

Michael Rafel Hosea Siregar

22/508531/PPA/06439

*Gaokao is the college entrance examination in China which often considered as the most difficult exam in the world. While existing studies have developed models to predict Gaokao admission scores, they still focused on the university level. This study aims to predict Gaokao admission scores at the major level, considering the level of competition among different study programs, and also determine the best competition model using elbow method.*

*In this research, the prediction consists of two stages, there are preliminary prediction and fine prediction. The preliminary prediction is conducted using several methods such as Gradient Boosted Decision Tree, Linear Regression, Multi Layer Perceptron, and Long Short-Term Memory (LSTM) models. After the preliminary prediction, each study program will be clustered using K-Means Clustering based on the admission scores from 2019 and the predicted admission scores for 2020 to obtain the average admission score for each majors in 2020 based on preliminary prediction results and the admission score of the major in the same position in the previous year defined as fine prediction.*

*This research shows that Gradient Boosted Decision Tree is the best method obtains a Mean Absolute Error of 5,1176, Mean Square Error of 58,0890, and a precision of 47,059% with a threshold of 3. In comparison, Long Short Term Memory obtains a Mean Square Error of 269,2584, a Mean Absolute Error of 14,4024, and a precision of 9,756% and a threshold of 3.*

***Keywords: Decision Tree, Linear Regression, Multi Layer Perceptron, Long Short Term Memory, Gaokao Score Prediction***