

# KLASIFIKASI KEPUASAN PELANGGAN *COFFEE SHOP* PATURUPA MENGUNAKAN *DECISION TREE*

Siti Hinggit

## ABSTRAK

*Coffee shop* saat ini sedang menjamur di Indonesia, dan juga persaingan saat ini pun juga memungkinkan *coffee shop* saat ini berlomba untuk meningkatkan kepuasan pelanggan. Paturupa merupakan salah satu coffee shop yang ada di Bekasi pun juga ingin mengetahui seberapa puas pelanggan yang ada di Paturupa dan faktor yang berpengaruh dalam kepuasan pelanggan. Dalam mewujudkan hal tersebut, diperlukan penggunaan data yang didapat dari kuesioner kepuasan pelanggan di Paturupa yang dibagikan melalui *google form*. Data yang didapatkan sebanyak 150 data berupa *class* “Ya” dan “Tidak” yang akan dibagi menjadi data latih dan data uji. Data latih dimanfaatkan untuk membangun model dengan algoritma *Decision Tree C4.5* yang memanfaatkan metode SMOTE dan NearMiss untuk mengatasi *imbalanced* data dan memperoleh nilai *entropy* dan *information gain* untuk menentukan akar node dan node lainnya. Hasil performa terbaik diperoleh penelitian ini untuk pembagian data latih 70% dan data uji 30% menggunakan metode NearMiss dengan akurasi sebesar 96%, *recall* sebesar 100%, *precision* sebesar 95%, dan *specificity* sebesar 75%. Variabel yang berpengaruh pada model adalah minuman dan makanan yang disajikan sesuai pada daftar menu.

**Kata Kunci:** Kepuasan Pelanggan, Data Mining, Klasifikasi, *Decision Tree C4.5*

# CLASSIFICATION OF CUSTOMER SATISFACTION COFFEE SHOP PATURUPA USING DECISION TREE

Siti Hinggit

## ABSTRACT

*Coffee shops are currently popular in Indonesia, and the current situation also allows coffee shops to compete to increase customer satisfaction. Paturupa one of coffee shop located in Bekasi want to know about how customer satisfaction in Paturupa and factor that is important for their customer satisfaction. For realizing, data are needed which are obtained from customer satisfaction questionnaires in Paturupa which are share through google forms. The data are obtained as many as 150 data with labels "Yes" and "No" which are divided into train data and test data. The train data are used to build a model with the Decision Tree C4.5 algorithm and use SMOTE and NearMiss for handling imbalanced data and obtained entropy and information gain values to determine the roots of nodes and other nodes. The best performance results are reached by this study with splitting data 70% for train data and 30% for test data that using NearMiss method, accuracy score is 96%, recall score is 100%, precision score is 95%, and specificity score is 75%.*

**Keywords:** *Customer Satisfaction, Data Mining, Classification, Decision Tree C4.5*