**Application of Decision Tree Algorithm for Classification and Predictive Modelling of UMKM Tahu Garing**

*Roma Sabet Manurung1, Kevin Junaidi2, Steven3, Steven4, Kristianti5*

*Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Internasional Batam, Sei Ladi, Jalan Gajah Mada, Baloi Permai, Sekupang, Batam, Kepulauan Riau, Indonesia 29442*

*1931104.Steven@uib.ac.id*

1. **Pendahuluan**

UMKM yang didirikan oleh Bapak Vincent ini diberikan nama Tahu Garing alias Tagar yang telah berdiri sejak bulan juli tahun 2019. Tagar (Tahu Garing) merupakan brand makanan ringan yang memiliki banyak cabang di Batam. Dari segi lokasi, Tagar memiliki stand yang didominasi warna jingga, sesuai dengan logo brand mereka. Terdapat etalase untuk memajang tahu-tahu yang baru digoreng, sehingga pelanggan juga bisa melihat. Tagar (Tahu Garing) mulai beroperasi pada pukul 11.00 WIB hingga pukul 19.00 WIB.

Masalah yang terjadi UMKM Tagar adalah dampak pandemi COVID-19. Pola pemasaran UMKM Tagar (Tahu Garing) pun berubah terlebih ketika diberlakukan *Social Distancing* dan Pembatasan Sosial Berskala Besar(PSBB). UMKM Bisnis Tagar (Tahu Garing) harus putar otak untuk bisa memasarkan produk atau jasa mereka ke konsumen, sebagai strategi *brand* bertahan di tengah pandemi COVID-19. UMKM Bisnis juga harus mengoptimalkan pemasaran *online* dan *digital branding* sebagai sarana komunikasi dengan target konsumen mereka.

Secara garis besar untuk kesuksesan pengembangan UMKM Tahu Tagar Garing dalam menciptakan solusi berbasis ML atau analisis data maka kami mendapat beberapa langkah yang bisa diperhatikan, yaitu [1] Memikirkan strategi akuisisi data dengan memanfaatkan kanal atau sumber daya yang ada. [2] Mengkomunikasikan nilai proposisi data untuk kebutuhan analisis. [3] Memastikan solusi dari algoritma yang dikembangkan berfokus untuk memberikan manfaat nyata bagi bisnis ataupun sasaran pengguna. Banyak pihak yang membutuhkan solusi berbasis *Machine Learning*. Bagi startup ataupun UMKM (pengembang) lainnya membangun set data ekslusif menjadi misi besar dalam menyukseskan sepak terjangnya. Tak mudah memang untuk mampu mendapatkan akuisisi pengguna, tapi setidaknya Internet mampu diandalkan untuk banyak hal.

Alasan pemakaian *Machine learning* dalam bisnis UMKM, yaitu [1] Investasi dalam teknologi ini dapat membantu UMKM. [2] Meningkatkan *agility* UMKM. [3] Membuka peluang pasar baru bagi UMKM. [4] Memahami pelanggan lebih baik.

1. **Metode Penelitian**
2. **Pengambilan *Dataset***

Kami mengambil *Dataset Customer* dari *website* Kaggle yang dimana *dataset* tersebut kami gunakan sebagai contoh apakah algoritma *Decision* *Tree* cocok untuk UMKM Tagar.

1. ***Classification and Predictive Modelling***

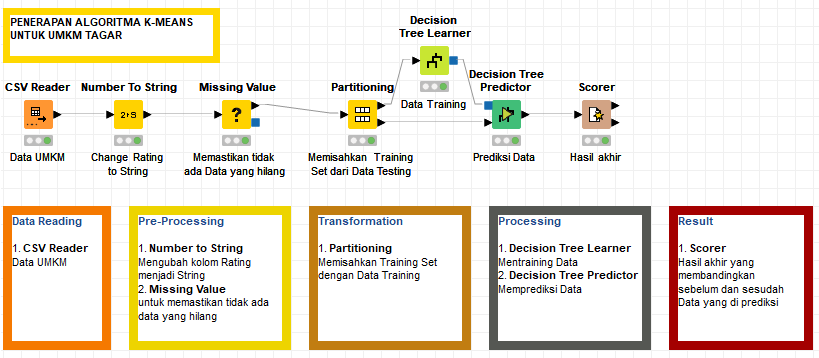
*Classification* dan *Predictive Modelling* merupakan metode utama dalam *Data Mining*. *Classification* adalah untuk menentukan sebuah *record data* baru ke salah satu dari beberapa *Class* atau *Category* yang telah didefenisikan sebelumnya.

1. **Decision Tree Algorithm**

Tujuan penggunaan *Decision Tree* adalah untuk membuat model pelatihan yang dapat digunakan untuk memprediksi *Class* atau nilai variabel target dengan **mempelajari aturan keputusan sederhana yang** disimpulkan dari data sebelumnya (*Training Data*).

1. **Hasil dan Pembahasan**

Dalam Penelitian ini kami menggunakan *Dataset Customer* dari *website* Kaggle dengan nama file “final\_rating.csv” sebagai contoh apakah algortima *Decision Tree* ini cocok untuk Tagar.

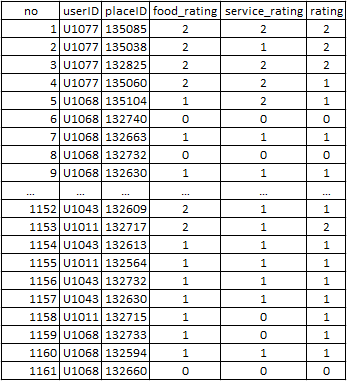


**Gambar 1. Model Penelitian KNIME**

1. ***Dataset* *Selection***

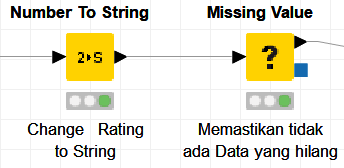
Kami menggunakan *dataset customer* bernama “final\_rating.csv” yang terdiri dari 1161 data yang diisi oleh *customer*.

**Table 1. Final Rating Data Table**



1. ***Dataset Pre-Processing***

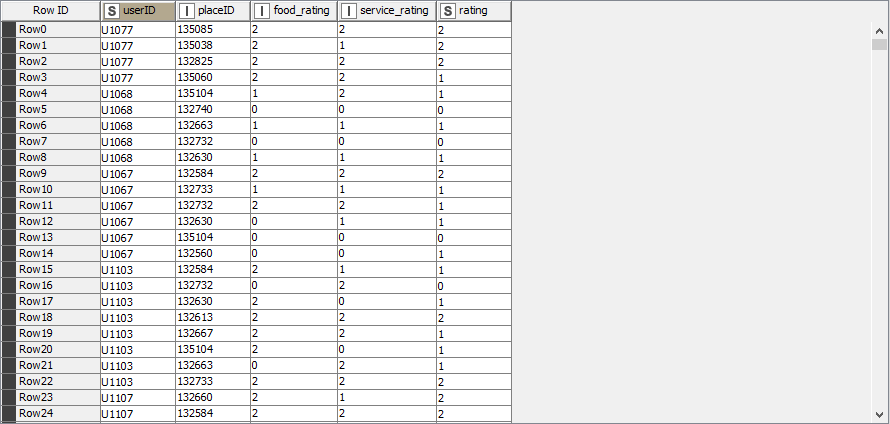
Pada tahap ini kami menggunakan node *Number to String* yang bertujuan untuk mengubah angka kolom *rating* menjadi *String* agar nanti dapat diprediksi. Kemudian, data dari *final rating* dibersihkan dengan node *Missing Value* yang dimana fungsinya yaitu mehilangkan data yang kosong.



**Gambar 2. Dataset Pre-Processing**

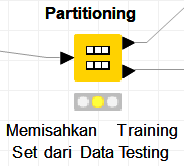
Berikut adalah hasil *dataset* dari pengubahan *string* dan dibersihkan:

**Table 2. Data Table of Missing Value**



1. ***Dataset Transformation***

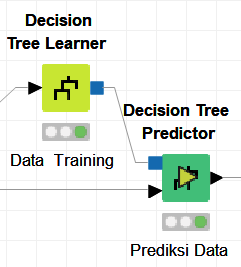
Pada tahap ini kami melakukan pemisahan dataset dengan perbandingan 9:1 menggunakan node *Partitioning* yang terdiri dari 90% *Data Training* dan 10% *Data Testing* dan memilih kolom rating sebagai classnya. *Data Training* akan diarahkan ke node *Decision Tree Learner* untuk proses *Classification*, sedangkan *Data Testing* diarahkan langsung ke node *Decision Tree Predictor*.



**Gambar 3. Partitioning Node**

1. ***Dataset Processing***

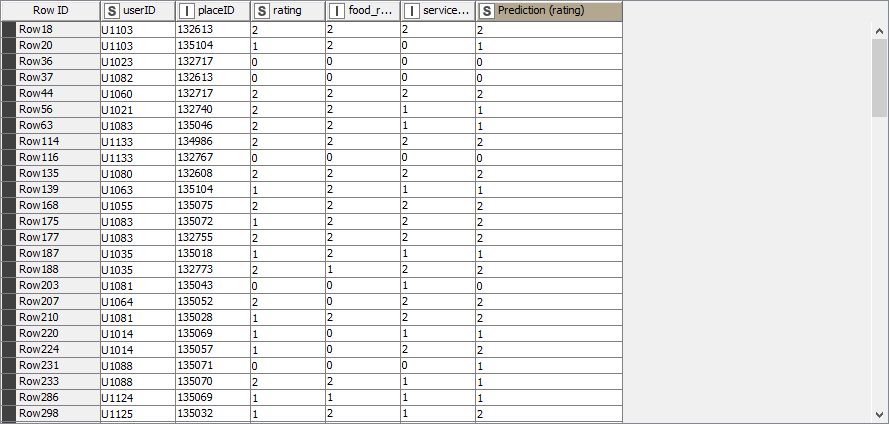
Tahap selanjutnya adalah *Classification* dengan menggunakan node *Decision Tree Learner* dan kemudian dilanjutkan ke node *Decision Tree Predictor*. Decision Tree Learner akan mengolah 90% data *final rating* kemudian setelah dipelajari oleh Decision Treenya kemudia dialurkan ke *Decision Tree Predictor* untuk memprediksi hasil kolom *rating* dari *dataset final rating*.

****

**Gambar 4. Dataset Processing**

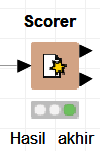
Berikut adalah hasil *dataset* dari *Decision Tree Predictor*:

**Table 3. Data Table of Decision Tree Predictor**



1. ***Result***

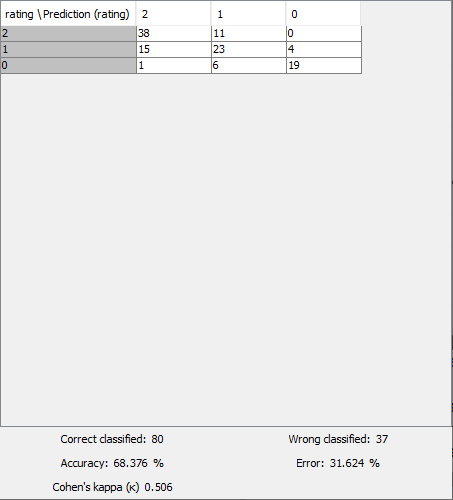
Tahap terakhir adalah menggunakan node *Scorer* untuk membandingkan kolom *rating* yang sudah diprediksi dengan kolom *rating* yang belum diprediksi dan melihat tingkat akurasi dari data tersebut.



**Gambar 5. Dataset Result**

Berikut adalah hasil node *Scorer*:

**Table 4. Data Table of Scorer**



Dari gambar table diatas dapat dilihat untuk kolom adalah rating yang sudah diprediksi dan baris adalah rating yang sebelum diprediksi. Data yang Benar Diklasifikasikan sebanyak 80 dan yang salah adalh 37 data, untuk tingkat *Accuracy* sebesar 68.376%

1. **Kesimpulan**

Dari hasil pengujian yang telah kami lakukan dengan *Dataset Customer* dari *website* Kaggle cocok digunakan untuk UMKM Tagar untuk memprediksi kekurangan dari Tagar apakah makan tersebut tidak enak atau pelayanan yang kurang.