

SKRIPSI

KAJIAN PREDIKSI *RATING* RESTORAN DI KOTA
BANDUNG DENGAN PENDEKATAN POHON KEPUTUSAN



Ambrosius Adrian Djunaidi

NPM: 6161801003

PROGRAM STUDI MATEMATIKA
FAKULTAS TEKNOLOGI INFORMASI DAN SAINS
UNIVERSITAS KATOLIK PARAHYANGAN
2022

FINAL PROJECT

**A STUDY OF RESTAURANTS RATING PREDICTION IN
BANDUNG CITY WITH DECISION TREE APPROACH**



Ambrosius Adrian Djunaidi

NPM: 6161801003

**DEPARTMENT OF MATHEMATICS
FACULTY OF INFORMATION TECHNOLOGY AND SCIENCES
PARAHYANGAN CATHOLIC UNIVERSITY
2022**

LEMBAR PENGESAHAN

KAJIAN PREDIKSI *RATING* RESTORAN DI KOTA BANDUNG DENGAN PENDEKATAN POHON KEPUTUSAN

Ambrosius Adrian Djunaidi

NPM: 6161801003

Bandung, 20 Januari 2022

Menyetujui,

Pembimbing 1



Liem Chin, M.Si.

Pembimbing 2



Agus Sukmana, M.Sc.

Ketua Tim Penguji



Dr. Livia Owen

Anggota Tim Penguji



Dr. Daniel Salim

Mengetahui,

Ketua Program Studi



Dr. Livia Owen

PERNYATAAN

Dengan ini saya yang bertandatangan di bawah ini menyatakan bahwa skripsi dengan judul:

KAJIAN PREDIKSI *RATING* RESTORAN DI KOTA BANDUNG DENGAN PENDEKATAN POHON KEPUTUSAN

adalah benar-benar karya saya sendiri, dan saya tidak melakukan penjiplakan atau pengutipan dengan cara-cara yang tidak sesuai dengan etika keilmuan yang berlaku dalam masyarakat keilmuan.

Atas pernyataan ini, saya siap menanggung segala risiko dan sanksi yang dijatuhkan kepada saya, apabila di kemudian hari ditemukan adanya pelanggaran terhadap etika keilmuan dalam karya saya, atau jika ada tuntutan formal atau non-formal dari pihak lain berkaitan dengan keaslian karya saya ini.

Dinyatakan di Bandung,
Tanggal 20 Januari 2022



Ambrosius Adrian Djunaidi
NPM: 6161801003

ABSTRAK

Rating restoran adalah ukuran kepuasan berdasarkan penilaian atau opini pelanggan restoran terhadap mutu dan fasilitas restoran. *Rating* yang tinggi (mendekati, lima dari lima) akan menarik orang untuk berkunjung ke suatu restoran. Oleh karena itu, dibangun sebuah model untuk membantu pengusaha restoran dalam memprediksi *rating* dari restoran yang akan dibuka atau untuk mempelajari pengaruh terhadap *rating* ketika mereka melakukan perubahan layanan. Model yang dibangun berdasarkan pohon keputusan dan menggunakan data *rating* dari ©Zomato™ Ltd untuk restoran di Kota Bandung dan sekitarnya. Kemudian, model pohon keputusan ini diperbaiki dengan memperhatikan parameter kedalaman maksimum dari pohon keputusan dan sampel minimal. Pemilihan kedua parameter ini dilakukan untuk menghasilkan pohon keputusan baru yang lebih baik ditinjau dari nilai koefisien determinasi. Selanjutnya, hasil prediksi *rating* restoran dari pohon keputusan ini beserta data lokasi restoran, dibuat visualisasi yang ramah pengguna berbentuk peta kecamatan dengan filter berupa fasilitas. Dari visualisasi ini, pengguna dapat memilih fasilitas untuk mendapatkan lokasi restoran dengan *rating* terbaik.

Kata-kata kunci: *Rating*, Prediksi, Pembelajaran Mesin, Pohon Keputusan, Visualisasi Data

ABSTRACT

A restaurant rating is a measure of satisfaction based on the rating or opinion of restaurant customers on the quality and facilities of the restaurant. A high rating (close to, five out of five) will attract people to visit the restaurant. Therefore, a model was built to assist restaurateurs in predicting the rating of a restaurant that will open or to study the effect on the rating when they make a service change. The model is built based on a decision tree and uses rating data from ©Zomato™ Ltd for restaurants in Bandung City and its surroundings. Then, this decision tree model is improved by taking into account the parameters of the maximum depth of the decision tree and the minimum sample. The selection of these two parameters is done to produce a new decision tree that is better in terms of the value of R^2 (coefficient of determination). Furthermore, the results of the restaurant rating predictions from this decision tree along with restaurant location data, a user-friendly visualization are made in the form of a sub-district map with a filter in the form of facilities. From this visualization, users can choose facilities to get the location of the restaurant with the best rating.

Keywords: Rating, Prediction, Machine Learning, Decision Tree, Data Visualization

Untuk Mama, Papa, dan Adik terbaik.

KATA PENGANTAR

Halo dunia! Puji syukur kepada Tuhan Yang Maha Esa karena atas berkat dan anugerah-Nya, penulis dapat menyelesaikan skripsi ini. Skripsi dengan judul “Kajian Prediksi *Rating* Restoran di Kota Bandung dengan Pendekatan Pohon Keputusan” ini disusun sebagai syarat untuk menyelesaikan studi sarjana di Program Studi Matematika, Fakultas Teknologi Informasi dan Sains, Universitas Katolik Parahyangan, Bandung.

Keberhasilan penyusunan skripsi ini tidak lepas dari hambatan dan kesulitan. Oleh karena itu, penulis mengucapkan terima kasih kepada semua pihak yang membantu penyusunan skripsi ini, terutama kepada:

- Kedua orang tua yang selalu memberikan dukungan, doa, dan nasihat untuk penulis, serta Adik penulis, Justin, yang selalu memberikan hiburan dan dukungan untuk penulis.
- Bapak Liem Chin, M.Si. selaku dosen wali, dosen pembimbing, dan dosen koordinator skripsi yang telah memberikan ilmu, arahan, dan saran dengan sabar baik selama penyusunan skripsi maupun perkuliahan.
- Bapak Agus Sukmana, M.Sc. selaku dosen pembimbing yang telah memberikan ilmu, arahan, dan saran dengan sabar baik selama penyusunan skripsi maupun perkuliahan.
- Ibu Dr. Livia Owen selaku dosen penguji yang telah memberikan kritik dan saran untuk skripsi ini, sehingga skripsi ini menjadi lebih baik.
- Bapak Dr. Daniel Salim selaku dosen penguji yang telah memberikan ilmu dan arahan selama perkuliahan serta kritik dan saran untuk skripsi ini, sehingga skripsi ini menjadi lebih baik.
- Seluruh dosen FTIS (Fakultas Teknologi Informasi dan Sains) yang telah memberikan ilmu, arahan, dan saran kepada penulis.
- Seluruh staf Tata Usaha FTIS yang telah banyak memberikan segala bentuk bantuan administratif kepada penulis selama perkuliahan.
- Yohanes Dimas Prasetyo yang memberikan dukungan dan hiburan serta sebagai teman yang menemani penulis selama berkuliah.
- Steven Sergio yang bertukar informasi dengan penulis selama penyusunan skripsi ini.
- Teman-teman HappyPadoru: JPJ (Jonathan Prasetyo Johan), Jojo (Jonathan Dinbara), Sergio, Topas Sagara, Malvin Hidayat, Novaldi, Rhandy, dan Yohanes Reinhart yang telah mendukung dan menghibur penulis.
- Teman-teman meatology: JPJ, Jojo, Sergio, Topas, dan Malvin yang menemani penulis berwisata kuliner selama berkuliah.
- Teman-teman apaniiiiiiiiiiii: Dimas, Novaldi, Orlin, dan Alwy yang memberikan dukungan dan hiburan kepada penulis.
- Teman-teman KO-MODO: Kathleen, Meliana, dan Christovito yang memberikan dukungan dan hiburan serta teman cerita bagi penulis.
- Cynthia Maria Henry dan Claudia Amaranantia yang memberikan dukungan dan hiburan serta teman cerita bagi penulis.
- Divisi Hubinsos (Hubungan Internal dan Sosial) HMPSMa 2020 yang memberikan ilmu, hiburan, dan pengalaman organisasi.
- Teman-teman seperjuangan skripsi yang memberikan dukungan selama penyusunan skripsi.
- Teman-teman Matematika angkatan 2017, 2018, 2019, dan 2020 yang tidak dapat disebutkan satu per satu.

- Seseorang yang ada di hati penulis namun tidak dapat disebutkan.

Penulis menyadari bahwa masih terdapat kekurangan dalam skripsi ini. Oleh karena itu, penulis menerima dengan terbuka semua kritik dan saran yang membangun dari pembaca untuk penyempurnaan skripsi ini. Semoga skripsi ini dapat bermanfaat dan menginspirasi pembaca.

Bandung, Januari 2022

Penulis

DAFTAR ISI

KATA PENGANTAR	xv
DAFTAR ISI	xvii
DAFTAR GAMBAR	xix
1 PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	1
1.3 Tujuan	2
1.4 Batasan Masalah	2
1.5 Sistematika Pembahasan	2
2 LANDASAN TEORI	3
3 METODE PENELITIAN	9
3.1 Persiapan Data	9
3.2 Pembentukan Model	11
3.3 Evaluasi Model	12
3.4 Visualisasi Hasil	12
4 HASIL DAN PEMBAHASAN	15
4.1 Data	15
4.2 Pohon Keputusan Awal	15
4.3 Evaluasi Pohon Keputusan	16
4.4 Visualisasi Prediksi <i>Rating</i>	21
5 KESIMPULAN DAN SARAN	23
5.1 Kesimpulan	23
5.2 Saran	23
DAFTAR REFERENSI	25
A DATA	27
B MODEL AWAL	29

DAFTAR GAMBAR

2.1	Contoh pohon keputusan	3
2.2	Alur pembentukan Gambar 2.1	4
2.3	Wilayah untuk Gambar 2.1 cabang ke-1 hingga 3	5
3.1	Metode penelitian	9
3.2	Pembentukan data	9
4.1	RSS kedalaman maksimum	16
4.2	R^2 sampel daun	16
4.3	Histogram selisih <i>rating</i> sebenarnya dan <i>rating</i> prediksi model awal	17
4.4	Histogram selisih <i>rating</i> sebenarnya dan <i>rating</i> prediksi model yang sudah dievaluasi	17
4.5	Diagram kotak garis model awal dan hasil evaluasi	18
4.6	Model pohon keputusan hasil evaluasi (1)	19
4.7	Model pohon keputusan hasil evaluasi (2)	20
4.8	Visualisasi prediksi <i>rating</i> restoran Kota Bandung	21
4.9	Visualisasi prediksi <i>rating</i> restoran Kota Bandung dengan filter	22
B.1	Model pohon keputusan awal (1)	29
B.2	Model pohon keputusan awal (2)	30
B.3	Model pohon keputusan awal (3)	31
B.4	Model pohon keputusan awal (4)	32
B.5	Model pohon keputusan awal (5)	33
B.6	Model pohon keputusan awal (6)	34
B.7	Model pohon keputusan awal (7)	35
B.8	Model pohon keputusan awal (8)	36
B.9	Model pohon keputusan awal (9)	37
B.10	Model pohon keputusan awal (10)	38
B.11	Model pohon keputusan awal (11)	39
B.12	Model pohon keputusan awal (12)	40
B.13	Model pohon keputusan awal (13)	41
B.14	Model pohon keputusan awal (14)	42
B.15	Model pohon keputusan awal (15)	43
B.16	Model pohon keputusan awal (16)	44

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Kota Bandung terkenal akan kulinernya, untuk itu banyak orang yang membuka usaha restoran karena merupakan lahan investasi yang baik. Kota Bandung juga terkenal karena beragamnya kuliner yang tersedia, dari berbagai negara dan daerah lokal. *Rating* adalah sebuah ukuran baik buruknya suatu hal yang diberikan oleh pelanggan. *Rating* merupakan ukuran yang paling menarik terutama untuk restoran, karena ketika orang sedang memilih makanan hal yang paling pertama dilihat, yaitu *rating* pada aplikasi kuliner. *Rating* juga dibentuk oleh orang-orang yang sudah pernah berkunjung, sehingga kelebihan maupun kekurangan sebuah restoran dapat diketahui terlebih dahulu.

Rating pada umumnya berkisar dari satu hingga lima. *Rating* yang tinggi (mendekati lima) dapat menarik minat orang baru untuk mencoba makanan di sebuah restoran atau kafe. Untuk itu dapat dibuat sebuah model untuk memprediksi hubungan sebuah restoran atau kafe dengan fasilitas-fasilitasnya agar memperoleh *rating* yang tinggi sehingga menarik banyak orang untuk mencobanya. Sebelumnya telah dilakukan penelitian untuk memprediksi tingkat sukses sebuah restoran secara global dengan data dari laman Zomato, untuk itu pada penelitian ini akan dilakukan di Indonesia khususnya Kota Bandung.

Pada penelitian yang telah dilakukan digunakan lima metode berbeda, yaitu *logistic regression*, *naive bayes*, pohon keputusan (*decision tree*), *random forest*, dan KNN [1]. Dari kelima metode diperoleh pohon keputusan yang terbaik dalam memprediksi tingkat kesuksesan. Untuk itu akan digunakan metode pohon keputusan untuk penelitian ini, karena pohon keputusan yang merupakan yang terbaik pada penelitian sebelumnya. Pada penelitian sebelumnya, setelah memperoleh model, hasil model tidak dipaparkan. Oleh karena itu, pada penelitian ini akan dibuat sebuah visualisasi yang dapat menentukan lokasi dengan prediksi *rating* terbaik.

Pada skripsi ini, *rating* restoran akan diprediksi dengan menggunakan model pohon keputusan. Hasil dari prediksi, kemudian akan divisualisasikan agar mempermudah mencari tempat dengan *rating* terbaik untuk sebuah restoran yang baru akan buka. Diharapkan dengan adanya penelitian ini pengusaha dapat terbantu untuk mencari lokasi yang strategis untuk membuka restoran utama maupun cabang restorannya.

1.2 Rumusan Masalah

Pada skripsi ini dibahas pembentukan model pohon keputusan untuk memprediksi *rating* restoran. Oleh karena itu, masalah yang akan dibahas dalam skripsi ini dirumuskan sebagai berikut.

1. Bagaimana membuat model yang menghubungkan antara fasilitas restoran dengan *rating* restoran menggunakan pohon keputusan?
2. Bagaimana mengukur performansi dari model pohon keputusan yang dihasilkan dan meningkatkan performannya?
3. Bagaimana memvisualisasikan hasil prediksi *rating* yang ramah pengguna dan memperoleh keputusan dari hasil visualisasi?

1.3 Tujuan

Berdasarkan rumusan masalah di atas, tujuan dalam skripsi ini adalah sebagai berikut.

1. Memperoleh model pohon keputusan antara fasilitas yang tersedia di restoran dengan *rating* restoran yang diperoleh.
2. Mengetahui performansi dari model yang dihasilkan dan proses meningkatkan performanya sehingga diperoleh model yang lebih akurat.
3. Mendapatkan visualisasikan hasil prediksi *rating* yang ramah pengguna.

1.4 Batasan Masalah

Skripsi ini hanya memprediksi *rating* restoran berdasarkan data yang diperoleh dari laman Zomato dan hanya berlokasi di kota bandung dan sekitarnya.

1.5 Sistematika Pembahasan

Sistematika pembahasan seminar ini terdiri dari 5 bab, yaitu:

1. Bab 1: Pendahuluan
Bab ini berisi latar belakang, rumusan masalah, tujuan, batasan masalah, dan sistematika pembahasan.
2. Bab 2: Landasan Teori
Bab ini berisi teori-teori yang mendukung dalam pengerjaan skripsi ini, yaitu pohon keputusan, pengertian dan cara pembuatan pohon keputusan sederhana.
3. Bab 3: Metode Penelitian
Bab ini berisi teknik pengambilan data, pembuatan model, evaluasi model, dan visualisasi hasil. Apabila setelah dievaluasi perlu dilakukan perbaikan, maka akan dilakukan perbaikan terlebih dahulu sebelum memvisualisasikan hasilnya.
4. Bab 4: Hasil dan Pembahasan
Bab ini berisi hasil dan pembahasan dari data untuk membuat model, model untuk memprediksi *rating* restoran serta visualisasi dari hasil prediksi.
5. Bab 5: Kesimpulan dan Saran
Bab ini berisi kesimpulan dan saran untuk pengembangan penelitian lebih lanjut dari skripsi ini.