

Rental Price Prediction of Boarding Houses in Batam City Using Linear Regression and Random Forest Algorithms

Yefta Christian^{1*}, Herman^{2*}, Jerry^{3*}

* Sistem Informasi, Universitas Internasional Batam

yefta@uib.ac.id¹, herman@uib.ac.id², khuJerry02@gmail.com³

Article Info

Article history:

Received 2023-10-30

Revised 2023-11-22

Accepted 2023-11-23

Keyword:

Price Prediction,
Boarding Houses,
Random Forest,
Linear Regression,
Scrum.

ABSTRACT

Boarding houses, commonly known as "kost," are residential places typically rented by individuals, serving a function similar to hotels, but with more affordable pricing. With the proliferation of boarding house businesses, residents and newcomers in Batam city face challenges in selecting suitable accommodation based on both price and amenities. Leveraging machine learning, a branch of artificial intelligence (AI), and incorporating various algorithms, a system can be developed to predict the rental prices of boarding houses. This helps individuals make informed decisions regarding the suitability of a boarding house based on their preferences and budget. The algorithms utilized in this study are Linear Regression and Random Forest. The modeling process resulted in R2 Scores, with Linear Regression achieving a score of 64%, while Random Forest outperformed with an impressive 99% R2 Score. Due to the higher R2 Score of Random Forest, this model was selected for the development of a website using the Scrum framework. The outcome of this research is a predictive pricing website for boarding houses, offering a valuable tool for residents and visitors in Batam when seeking to rent or lease a boarding house.



This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

I. PENDAHULUAN

Rumah atau tempat tinggal termasuk kedalam salah satu kebutuhan pokok yang sangat dibutuhkan oleh setiap manusia yaitu papan [1]. Dalam hal merantau, jika berasal dari kalangan menengah keatas, pasti akan menyewa sebuah apartemen maupun membeli sebuah rumah untuk tempat tinggal mereka. Sedangkan jika berasal dari kalangan menengah kebawah pasti tidak akan sanggup untuk membeli sebuah rumah maupun menyewa apartemen. Dikarenakan harga untuk membeli sebuah rumah maupun menyewa apartemen sangat mahal. Jadi, kalangan menengah kebawah pasti akan memilih untuk menyewa sebuah rumah kost maupun kamar kost [2].

Kost adalah sebuah tempat tinggal yang disewakan kepada orang-orang dengan memiliki fasilitas tertentu yang harganya sangat bervariasi dan juga lebih murah dibandingkan dengan hotel maupun apartemen [3]. Kini kost sangat banyak ditemui dimanapun, karena bisnis kost sekarang sudah menjamur kemana-mana, termasuk untuk kota Batam [4]. Banyak para perantau yang datang ke Batam baik untuk bekerja maupun

untuk menuntut ilmu. Untuk memilih kost pasti akan membutuhkan beberapa fasilitas yang sudah tersedia didalamnya untuk membantu aktifitas di dalamnya. Akan tetapi biasanya jika sebuah kost tersebut memiliki fasilitas yang lengkap maka harga dari sebuah kost akan sangat berpengaruh [5]. Bukan dari fasilitas saja akan tetapi letak yang strategis juga akan sangat mempengaruhi harga dari sebuah kost [6]. Karena dengan banyaknya bisnis kost dengan fasilitas dan harga yang berbeda maka para perantau akan mengalami kebingungan ketika ingin menyewa sebuah kost di Batam.

Pada akhir dekade ini teknologi sudah sangat akrab terdengar dan digunakan oleh manusia. Contohnya seperti mobil yang mempunyai autopilot, asisten virtual dirumah dan lain-lain. Teknologi ini disebut juga dengan Artificial Intelligence (kecerdasan buatan). Dengan adanya teknologi kecerdasan buatan (AI) maka kebutuhan manusia saat ini sudah sangat dimudahkan dengan mesin maupun teknologi. Machine learning merupakan salah satu dari cabang AI. Machine learning dikembangkan agar setiap mesin ini bisa bekerja dengan sendirinya tanpa bantuan dari pengguna [7].

Machine learning didasarkan oleh ilmu statistika, data mining, matematika dan lain-lain. Teknik-teknik machine learning memiliki banyak tipe penggunaannya, semuanya bisa digunakan berdasarkan kebutuhan dari bidang penggunaannya [8]. Dengan ini, maka machine learning bisa dijadikan sebagai alat untuk bisa membuat sebuah sistem yang bisa memprediksi harga yang sesuai dengan fasilitas dan juga letak [9] dari sebuah kost di kota Batam. Dengan menambahkan algoritma supervised learning yaitu linear regression dan random forest kedalam machine learning maka nantinya sistem ini akan bisa membuat sebuah prediksi dari sebuah harga kost tersebut. Dengan begitu maka akan bisa membantu para perantau atau masyarakat yang ingin menyewa kost di kota Batam.

Pada penelitian [10] yang berfokus kepada pembuatan sistem yang bisa memprediksi harga sewa kamar kost dengan menggunakan algoritma multiple linear regression. Tujuan dari penelitian ini adalah bisa membuat sebuah aplikasi berbasis website yang bisa membantu para anak kost memprediksi harga dari sebuah kost didaerah yang dekat dengan Universitas Kristen Maranatha. Metode yang di pakai dalam penelitian ini adalah pendekatan kuantitatif untuk pengumpulan data dan algoritma multiple linear regression untuk melakukan training dan testing pada data sehingga bisa memprediksi harga kost. Hasil dari penelitian ini adalah sistem yang dikembangkan dapat membantu calon penghuni kamar kost atau pebisnis kamar kost dalam mengetahui kisaran harga sewa kamar kost dan hasil ini mendapatkan akurasi 70%.

Kemudian pada penelitian [11] ini berfokus kepada analisis tren kepemilikan rumah dan juga prediksi dari harga rumah tersebut. Tujuan dari penelitian ini adalah untuk mengetahui faktor-faktor yang mempengaruhi dari harga sebuah rumah dan juga untuk melihat pola penyebaran pemodelan sederhana tentang simulasi harga rumah berdasarkan keinginan masyarakat. Metode yang dipakai pada penelitian ini adalah tahap analisis untuk pengumpulan data dan juga pemodelan data dengan menggunakan algoritma K-Means dan juga random forest, kemudian tahap deployment dengan menggunakan framework flask kemudian akan dikembangkan sebuah aplikasi sederhana. Hasil dari penelitian ini adalah sebuah website sederhana yang dibuat hanya dengan HTML dan kemudian bisa menganalisa dari penyebaran perumahan, prediksi harga rumah, dan juga prediksi harga rumah dengan pemodelan yang diinginkan oleh masyarakat.

Selanjutnya pada penelitian [12], mereka berfokus kepada pembuatan sebuah sistem yang bisa melakukan prediksi untuk penjualan sebuah apartemen. Tujuan penelitian ini adalah untuk membuat sebuah sistem yang mana nantinya akan bisa memprediksi berapa banyak keuntungan atau kerugian sebuah project apartemen yang nantinya akan dijual. Penelitian ini menggunakan Monte Carlo Method untuk mempelajari bagaimana data nantinya akan di olah, dan kemudian untuk pengembangan sebuah websitenya akan menggunakan kerangka kerja SCRUM yang diantara lain

yaitu product backlog, sprint backlog, sprint planning, sprint meeting, daily scrum, sprint review, sprint retrospectives. Hasil dari penelitian ini adalah sebuah website yang bisa memprediksikan hasil dari penjualan dari apartemen baik itu kerugian maupun keuntungannya.

Pada penelitian [13] mereka berfokus untuk menyelesaikan permasalahan pada lama waktu dalam memberikan sebuah persetujuan ketika para karyawan melakukan penginputan disaat ingin melakukan pengajuan pembelian pada PT. Sumber Alfaria Trijaya, Tbk karena masih memakai form manual. Tujuan dari penelitian ini adalah membuat sebuah aplikasi sistem manajemen workflow yang bisa meningkatkan keefektifitas terhadap waktu serta bisa membantu karyawan baik dalam melakukan purchasing request dan juga approval request. Metode yang digunakan dalam penelitian ini adalah metode wawancara untuk mendapatkan sekumpulan informasi yang akan berguna dalam pengembangan sebuah aplikasi. Sedangkan untuk pengembangan sistem memakai metode Rapid Application Development (RAD) serta juga memanfaatkan framework flask untuk pembangunan backend agar lebih mempermudah para developer backend dan juga memanfaatkan Jinja Template untuk membangun UI oleh para developer frontend. Hasil dari penelitian ini adalah sebuah sistem purchasing approval request yang bisa mempermudah para karyawan karena mempunyai fitur purchasing request, approval request dan juga pengelolaan data request yang masuk.

Dan kemudian yang terakhir pada penelitian [14] berfokus kepada sistem pendukung keputusan terhadap pemilihan rumah kost berdasarkan harga untuk mahasiswa IT Telkom Purwokerto. Tujuan dari penelitian ini adalah untuk bisa membuat sebuah sistem yang bisa membantu mahasiswa dalam memilih kamar kost yang sesuai dengan kriteria yang diinginkan berdasarkan harga. Metode yang dipakai didalam penelitian ini adalah studi literatur dan juga observasi untuk mendapatkan data yang dibutuhkan dan selanjutnya akan dikembangkan menggunakan algoritma fuzzy tsukamoto untuk bisa membuat sebuah sistem yang bisa memprediksi harga kamar kost. Hasil dari penelitian ini adalah sebuah website yang bisa memprediksikan harga dari sebuah rumah kost yang diinginkan dari seseorang dengan hanya memasukkan atau mencentang pilihan yang ada didalam website tersebut dan nantinya akan muncul harga dari apa yang kita prediksi.

Berdasarkan tinjauan pustaka diatas, maka penelitian ini menerapkan machine learning untuk memprediksi harga dari sebuah kost dengan menggunakan algoritma Linear Regression [10] dan Random Forest [11] dengan mengklasifikasi berdasarkan kecamatan, ukuran kamar, listrik dan air, dan juga fasilitas yang tersedia. Menggunakan kerangka kerja Scrum untuk pengembangan website [12] dan juga menggunakan framework flask untuk bagian backend [13], setelah itu akan dihosting sehingga bisa digunakan masyarakat [14]. Yang membedakan penelitian kali ini dengan penelitian sebelumnya adalah dimana didalam penelitian ini menggunakan 2 algoritma yang berbeda

sehingga bisa menentukan algoritma yang lebih cocok untuk dijadikan model prediksi harga kost. Tujuan dari penelitian ini yaitu untuk mengklasifikasikan nilai harga kost berdasarkan fasilitas-fasilitas sehingga menghasilkan sebuah nilai prediksi yang kemudian akan di kembangkan menjadi sebuah website yang bisa digunakan oleh para perantau dan masyarakat di kota Batam.

II. METODE PENELITIAN

Alur yang penulis gunakan dalam penelitian ini adalah sebagai berikut.



Gambar 1. Alur Penelitian

Penelitian ini dimulai dengan pengumpulan data yang akan di kumpulkan dari beberapa kecamatan yang ada di batam dengan menggunakan kuesioner. Ketika data sudah terkumpul kemudian dilanjutkan dengan pengolahan data yang akan dilakukan pemodelan dengan menggunakan algoritma linear regression dan juga random forest agar bisa melakukan prediksi. Dari pemodelan ini akan menghasilkan sebuah R2 Score, yang dimana R2 Score dari model yang tertinggi akan digunakan untuk pengembangan website. Setelah memutuskan model yang akan di pakai, kemudian akan di lanjutkan dengan pengembangan website prediksi menggunakan kerangka kerja Scrum, dan ketika website nya sudah selesai maka akan di lakukan pengujian terhadap website dengan menggunakan *blackbox testing* agar tidak terjadi error yang tidak disengaja selama pengembangan. Dan kemudian akan dilakukan hosting agar para user bisa menggunakan website tersebut.

A. Teknik Pengumpulan Data

Pada tahapan pengumpulan data ini akan menggunakan metode pendekatan kuantitatif dimana penulis akan menyebarkan kuesioner survey dengan menggunakan google form. Selain kuantitatif penulis juga menggunakan pendekatan survey lapangan, dimana akan terjun langsung ke lapangan untuk menanyakan data tentang kost. Kuesioner survey ini akan disebar ke seluruh para pebisnis kost dan juga para pengguna kost yang ada di kota Batam. Adapun beberapa pertanyaan yang akan dimasukkan kedalam kuesioner ditunjukkan kedalam tabel I dibawah ini.

TABEL I
KETERANGAN KATEGORI SURVEY

No	Nama	Kategori	Keterangan
1	Kecamatan	0 = Batam Kota 1 = Batu Aji 2 = Batu Ampar 3 = Bengkong 4 = Lubuk Baja 5 = Nongsa 6 = Sagulung 7 = Sekupang	Kecamatan yang dipilih untuk penelitian ini karena dimana data yang akan didapatkan cukup untuk dilakukan penelitian
2	Ukuran Kamar	0 = 2 x 2 1 = 2 x 3 2 = 3 x 3 3 = 3 x 4 4 = 4 x 4	Pembagian ukuran kamar
3	Listrik Air	0 = Air Saja 1 = Termasuk Listrik	Fasilitas yang disediakan pemilik kost
4	Fasilitas Dalam Kamar	1. AC 2. Air Panas 3. WC Dalam 4. Kasur 11. Tidak Ada	Fasilitas yang didapatkan di dalam kamar kost
5	Fasilitas Luar Kamar	1. Dapur 2. WC 3. Kulkas 4. Laundry 9. Wifi	Fasilitas umum yang di dapatkan dari kost
6	Harga Sewa	0 = Rp.500.000 - Rp.750.000 1 = Rp.751.000 - Rp.1.000.000 2 = Rp.1.001.000 - Rp.1.250.000 3 = Rp.1.251.000 - Rp.1.500.000 7 = Rp.2.251.000 - Rp.2.500.000	Harga yang dibayar untuk menyewa kamar kost

Berdasarkan tabel diatas terlihat ada 6 kategori yang akan penulis masukkan kedalam kuesioner dan kemudian disebar nantinya untuk mendapatkan data. Diantara 6 kategori tersebut terdapat 24 atribut dan ada 1 atribut yang dipengaruhi oleh atribut lain yaitu atribut Harga Sewa.

B. Pengolahan Data

Pada pengolahan data ini akan dilakukan menggunakan perangkat lunak jupyter notebook version 6.4.5 dan juga Microsoft Excel Office 2016. Ada beberapa tahap yang akan dilakukan didalam pengolahan data adalah sebagai berikut.

1) *Pembersihan Data*

Pada data yang telah didapatkan, penulis akan melakukan pembersihan terhadap data yang kosong. Pembersihan data ini bisa dilakukan dengan 2 cara. Cara pertama yaitu dengan cara menghapus baris data yang mempunyai kolom kosong di file Microsoft Excel. Cara kedua yaitu dengan menggunakan code di bahasa pemrograman python.

2) *Pengubahan Bentuk Data (Transformasi Data)*

Data yang telah dibersihkan tadi kemudian akan di lakukan perubahan bentuk data. Data yang berbentuk kata akan diubah menjadi data yang bernilai numerik. Pada perubahan bentuk data ini akan dilakukan dengan menggunakan encoder mapping. Sehingga data bisa dipahami oleh pembelajaran mesin.

3) *Training dan Testing Data*

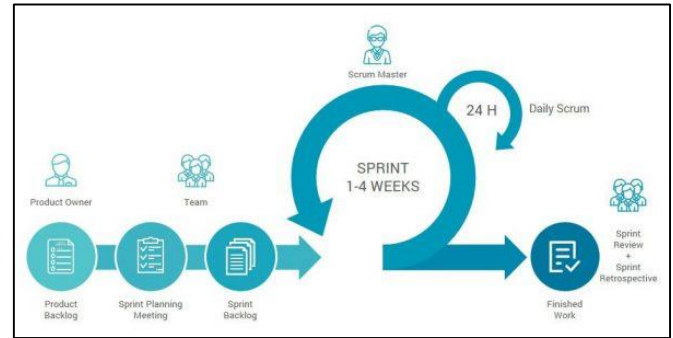
Setelah data melewati tahapan label encoder maka data akan dibagi menjadi dua bagian yaitu menjadi data training dan data testing. Setelah dilakukan pembagian data ini maka langkah selanjutnya adalah melakukan pemodelan dengan menggunakan algoritma linear regression dan juga random forest. Setelah melakukan pemodelan kemudian akan dilanjutkan dengan melakukan training dengan menggunakan data training. Setelah dilakukannya training maka langkah selanjutnya adalah melakukan test dengan menggunakan data testing sehingga bisa menghasilkan harga prediksi.

4) *Evaluasi*

Pada tahap terakhir ini, akan dilakukan evaluasi terhadap data yang telah bisa melakukan prediksi diatas. Evaluasi ini akan menghasilkan sebuah nilai MAE, MSE, dan RMSE dan juga R2 Score. Jika nilai R2 Score tinggi maka harga yang keluar dari hasil prediksi ini juga sangat bagus karena tingkat keakuratannya sangat tinggi. Dan sebaliknya jika nilai MAE, MSE dan RMSE semakin rendah maka semakin baik prediksi yang akan di dapatkan [15]. Nilai R2 Score juga merupakan tingkat akurasi dari prediksi, makanya jika semakin besar nilai R2 Score maka akan lebih besar nilai dari akurasi nya, dan begitu juga sebaliknya. Setelah mendapatkan hasil tertinggi maka model dari algoritma tersebut di ekspor menjadi bentuk file pkl untuk digunakan dalam pengembangan website.

C. Pengembangan

Pada tahap pengembangan, penulis akan menggunakan kerangka kerja Scrum untuk membuat sebuah website prediksi harga sewa kost.



Gambar 2. Alur Kerangka Kerja SCRUM

Kerangka kerja Scrum memiliki 3 fase dalam tahap pengembangannya, antara lain adalah sebagai berikut [16] .

1) *Fase Pre-Game*

Pada fase ini akan dimulai dengan menyusun perencanaan proyek sistem informasi yang ingin dirancang. Dimulai dengan membuat daftar kebutuhan serta penyusunan aktivitas yang akan dikerjakan nanti. Dalam fase ini akan menghasilkan luaran yaitu product backlog.

2) *Fase Game*

Dalam fase ini, pengembangan sebuah produk akan dimulai. Fase ini disebut juga dengan sprint atau jantung dari SCRUM. Adapun beberapa kegiatan yang akan dilakukan dalam fase ini adalah:

- **Sprint Planning**, merupakan awal dari kegiatan yang dilakukan. Dimana di dalam kegiatan ini akan menentukan pembagian tugas, estimasi waktu dan juga nilai kepentingan dari product backlog yang akan menghasilkan Sprint backlog.
- **Daily Scrum Meeting**, merupakan kegiatan dimana akan melakukan laporan hasil pekerjaan yang telah dilakukan. Pekerjaan disini dilakukan sesuai dengan sprint backlog. Dan disini lah model yang telah di ekspor ke bentuk pkl tadi digunakan untuk pembuatan website. Model tersebut dikembangkan dengan framework flask.
- **Sprint Review**, dimana ini merupakan kegiatan yang dilakukan di akhir daily scrum meeting. Dikegiatan ini akan melakukan peninjauan ulang atas hasil pekerjaan yang telah dilakukan.
- **Sprint Retrospectives**, merupakan kegiatan dimana akan melakukan pembahasan untuk pengembangan apalagi yang dibutuhkan didalam sebuah produk tersebut dan langsung melakukan perencanaannya di hari itu.

3) *Fase Post-game*

Pada fase ini, website yang telah dikembangkan kemudian di hosting sehingga para perantau dan masyarakat kota Batam bisa menggunakan website tersebut.

Pada tahap ini penulis akan menggunakan sebuah perangkat lunak untuk menulis editor code yaitu pada Visual Studio Code version 1.73.0 dan juga Google Chrome.

D. Pengujian

Pada fase pengujian penulis akan menggunakan blackbox testing. Pada pengujian blackbox ini setelah website telah selesai di kembangkan, maka penulis akan menguji dari website tersebut dengan memasukkan nilai nilai yang sama seperti pada uji prediksi pada model nya secara langsung. Dan kemudian akan dilihat dari hasilnya yang dari model dan juga dari website, jika tidak sama maka telah terjadi kesalahan dalam melakukan pengkodean flask.

III. HASIL DAN PEMBAHASAN

A. Pengumpulan Data

Setelah melakukan pengumpulan data dengan menyebarkan kuesioner dengan google form dan juga survey lapangan, penulis mendapatkan sebanyak 1.504 data dari beberapa kost yang tersebar di kota Batam. Dari data yang telah didapatkan maka penulis bisa menyimpulkan seperti dibawah ini:

1. Ada 24 atribut yang memengaruhi dari sebuah harga yaitu kecamatan, ukuran, listrik air, dan 11 atribut dari fasilitas kamar dan 10 atribut dari fasilitas bersama.
2. Terdapat sebanyak 1.054 data dari kost yang memiliki 8 kecamatan yang berbeda beda, 4 jenis ukuran kamar dan juga 8 jenis harga yang berbeda beda.

B. Pengolahan Data

Hasil dari pengolahan data adalah sebagai berikut.

1) Pembersihan Data

Pembersihan data pada tahap pembersihan data kali ini penulis menggunakan Microsoft Excel. Dengan mencari data yang kosong dan juga data yang memiliki perbedaan tetapi hanya sedikit maka data tersebut akan di hapus agar tidak mengganggu dalam proses pemodelan. Setelah melakukan pembersihan data di excel penulis juga memastikan lagi tidak adanya data yang kosong ketika ingin melakukan pemodelan di jupyter notebook seperti pada gambar dibawah berikut.

```
df.isna().mean()*100
```

Gambar 3. Code Melihat Data Kosong

2) Pengubahan Bentuk Data (Transformasi Data)

Pada tahap pengubahan bentuk data, penulis menggunakan Encoder Mapping untuk pengubahan bentuk data type object ke bentuk angka. Sehingga mesin lebih bisa memahami dan mengolah data dan juga penggunaan encoder mapping ini akan lebih memudahkan penulis dalam mengimplementasikan flask nantinya. Penggunaan encoder mapping seperti pada gambar di bawah ini.

```
def get_integer_mapping(le):
    res = {}
```

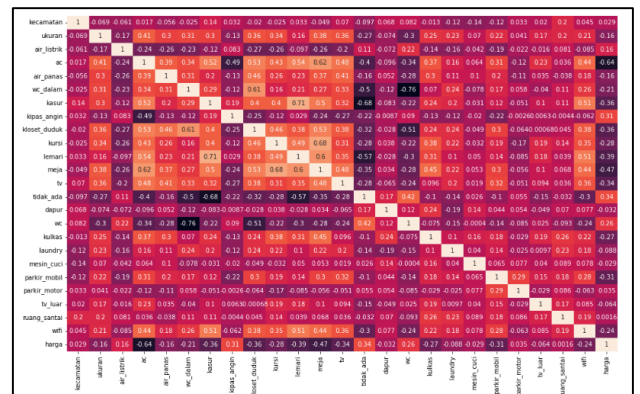
```
for cl in le.classes :
    res.update({cl:le.transform([cl])[0]
})

return res

ENCODER_MAPPING = dict()
for key in df.keys():
    encoder = LabelEncoder()
    ENCODER_MAPPING[key] = dict()
    for key2, val in
get_integer_mapping(encoder.fit(df[key])).items():
        ENCODER_MAPPING[key][key2] = val
        ENCODER_MAPPING[key][val] = key2
        ENCODER_MAPPING[key]['encoder'] = encoder
df[key] = list(map(lambda x:
ENCODER_MAPPING[key][x], df[key]))
```

Gambar 4. Penggunaan Code Encoder Mapping

Setelah data diubah menggunakan encoder, penulis melanjutkan untuk melakukan analisis dengan menggunakan subplot dan hasilnya seperti dibawah ini.



Gambar 5. Heatmap Subplot

Dari gambar diatas bisa disimpulkan bahwa dari 11 fasilitas yang ada didalam kamar, atribut kipas_angin dan atribut tidak_ada yang sangat mempengaruhi sebuah harga kost. Sedangkan dari 10 fasilitas bersama hanya atribut wc lah yang sangat mempengaruhi dalam sebuah harga kost.

3) Training Testing Data dan Evaluasi Data

Setelah melakukan analisa pada data maka penulis selanjutnya membagikan data untuk training dan juga testing nantinya. Data dibagi menjadi 90% data training dan 10% data testing, dimana berarti data training memiliki 1.354 data dan data testing 150 data.

```
x_train, x_test, y_train, y_test =
train_test_split(x, y, test_size=0.1)
```

Gambar 6. Pembagian Data Testing

Setelah melakukan pembagian data, selanjutnya pemodelan algoritma linear regression dan random forest akan diterapkan.

```
lin_reg = LinearRegression()
lin_reg.fit(x_train, y_train)
predictions = lin_reg.predict(x_test)
```

Gambar 7. Pemodelan Linear Regression

```
random_forest =
RandomForestClassifier(n_estimators=100)
random_forest.fit(x_train, y_train)
predictions = random_forest.predict(x_test)
```

Gambar 8. Penerapan Random Forest

Setelah melakukan penerapan menggunakan algoritma maka selanjutnya adalah evaluasi untuk melihat hasil dari algoritma yang telah di terapkan. Evaluasi ini menggunakan nilai MAE, MSE, RMSE dan juga R2 Score.

```
MAE: 1.106584465630459
MSE: 2.0274342102870744
RMSE: 1.4238799845096055
R2 Score: 0.6440915308135167
```

Gambar 9. Nilai dari Algoritma Linear Regression

```
MAE: 0.019867549668874173
MSE: 0.019867549668874173
RMSE: 0.14095229572048187
R2 Score: 0.9965123261937391
```

Gambar 10. Nilai dari Algoritma Random Forest

Dari gambar diatas bisa lihat nilai dari linear regression dari MAE nya 1,11, MSE nya 2.03, RMSE 1,42 dan R2 Score nya 64%. Sedangkan dari Random Forest MAE 0.02, MSE 0.02, RMSE 0.14 dan R2 Score nya 99%. Nilai R2 Score ini adalah nilai akurasi, maka algoritma yang terbaik untuk memprediksi harga kost adalah menggunakan random forest dengan 99% tingkat akurasi. Setelah selesai melakukan evaluasi maka model akan di ekspor menjadi tipe pkl menggunakan library dari pickle.

```
pickle.dump(random_forest, open("model.pkl",
"wb"))
pickle.dump(ENCODER_MAPPING,
open("encoder_mapping.pkl", "wb"))
```

Gambar 11. Ekspor model

C. Pengembangan

Pada tahap pengembangan ini, penulis menggunakan kerangka kerja scrum untuk pembuatan website nya. Tahap pertama dari scrum adalah fase pre-game dimana didalam fase ini penulis menentukan product backlog yang nantinya akan digunakan untuk list kegiatan pengembangan.

TABEL II
PRODUCT BACKLOG

No	User Story	Skala Prioritas
1	Pembuatan banner awal website	Tinggi
2	Pembuatan form prediksi	Tinggi

3	Pengerjaan backend menggunakan flask	Tinggi
4	Melakukan hosting	Sedang

Tabel diatas merupakan tabel dari product backlog dimana nantinya tabel ini akan digunakan untuk penyusunan dari sprint planning.

TABEL III
SPRINT BACKLOG

No	User Story	Prioritas	Waktu (Hari)
1	Pembuatan banner awal website	1	1
2	Pembuatan form prediksi	1	1
3	Pengerjaan backend menggunakan flask	2	4
4	Melakukan hosting	3	2

Tabel diatas merupakan tabel sprint backlog yang sudah terisi dengan estimasi waktu penyelesaian untuk pengerjaan project. Sprint backlog ini dibuat di fase sprint planning. Setelah sprint backlog selesai dibuat maka penulis akan memulai pengerjaan project. Pada awal pembuatan website penulis membuat banner dan kemudian akan di lanjutkan dengan pembuatan form untuk prediksi.

```
app = Flask( name )
model = pickle.load(open('model.pkl', 'rb'))
ENCODER_MAPPING =
pickle.load(open('encoder_mapping.pkl', 'rb'))

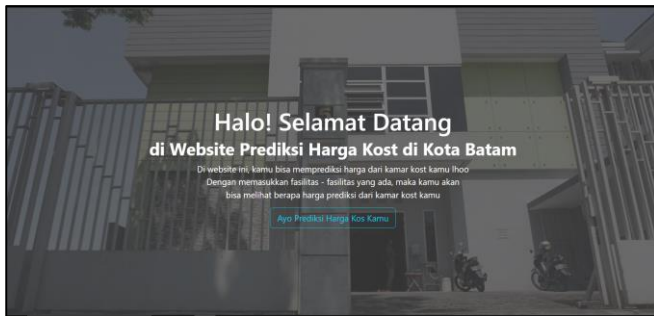
@app.route('/', methods=['GET'])
def home():
    return render_template('index.html')

@app.route('/predict', methods=['POST'])
def predict():
    colnames =
['kecamatan', 'ukuran', 'air listrik', 'ac', 'air pa
nas', 'wc_dalam', 'kasur', 'kipas angin', 'kloset_du
duk', 'kursi', 'lemari', 'meja', 'tv', 'tidak ada', 'd
apur', 'wc', 'kulkas', 'laundry', 'mesin cuci', 'park
ir mobil', 'parkir motor', 'tv luar', 'ruang santai
', 'wifi']
    info_for_prediction = []
    for key in colnames:
        if request.form.get(key, None):
            info_for_prediction.append(ENCODER_M
APPING[key] ['Yes' if request.form[key] ==
'check' else request.form[key]])
        else:
            info_for_prediction.append(ENCODER_M
APPING[key] ['No'])
    prediction =
model.predict([info_for_prediction])
    return jsonify({'harga':
ENCODER_MAPPING['harga'][prediction[0]] })
if __name__ == '__main__':
    app.run(port=3000, debug=True)
```

Gambar 12. Pengkodean Flask

Selanjutnya jika pengembangan design dari halaman website telah selesai maka akan dilanjutkan tahap penghubungan model prediksi dengan website menggunakan framework dari flask seperti pada gambar 12. diatas. Pada tahap ini penulis juga menggunakan encoder mapping untuk merubah data yang nantinya akan di input dari form sehingga

mesin bisa melakukan prediksi dengan model yang telah dibuatkan sebelumnya.

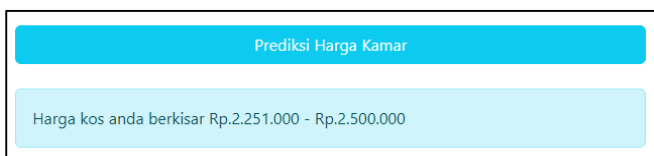


Gambar 13. Banner Depan Halaman Website

Gambar 14. Form Prediksi Harga Kost

Setelah tahap flask selesai maka akan dilanjutkan pada tahap hosting. Tahap ini dibutuhkan agar semua orang bisa menggunakan website ini. Website ini di hosting di pythonanywhere yang mana ini merupakan website layanan hosting yang memiliki akses dari gratis sampai berbayar.

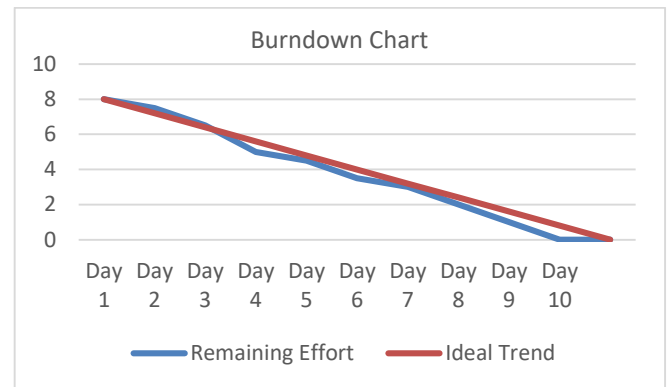
Pada Gambar 13 merupakan banner dari website prediksi, jika baru membuka website prediksi maka akan muncul banner dan jika user mengklik tombol “Ayo Prediksi Harga Kost Kamu” atau scroll kebawah maka akan muncul form seperti pada gambar 14. Disana user bisa mengisi data sesuai dengan fasilitas yang dikamar kost user atau bisa sesuai dengan yang inginkan. Seperti Kecamatan, Ukuran Kamar, Termasuk Air Listrik, Fasilitas Kamar, dan Fasilitas Luar Kamar. Kemudian jika menekan tombol “Prediksi Harga kamar” maka akan muncul prediksi dari harga kamar seperti pada gambar dibawah ini.



Gambar 15. UI Hasil Prediksi

Dan ketika hosting telah selesai maka tahapan dari scrum sudah berakhir. Website ini bisa digunakan oleh seluruh para perantau dan juga masyarakat yang ada di kota Batam baik dari untuk menyewa sebuah kost maupun bagi para pemilik kost yang ingin menyewakan kost nya. Pada sprint planning

diatas menghasilkan sebuah chart yang disebut burndown chart, yang dimana fungsi dari chart ini adalah untuk melihat apakah pekerjaan yang telah di list pada product backlog tadinya berjalan dengan sesuai jadwal. Pada chart dibawah bisa disimpulkan bahwa sprint planning pada penelitian ini berjalan hampir sempurna sesuai dengan jadwal.



Gambar 16. Burndown Chart Scrum

D. Pengujian

Tahap pengujian ini menggunakan metode blackbox testing, dengan memasukkan data yang sama di model prediksi dan juga di website prediksi dan hasilnya adalah sebagai berikut.

TABEL IV
PENGUJIAN BLACKBOX TESTING

No	Pengujian	Hasil Prediksi	Kesimpulan
1a	Lubuk Baja, 3 x 4, Termasuk Listrik, AC, WC Dalam, Kasur, Kloset Duduk, Kursi, Lemari, Meja, TV, Parkir Mobil, Parkir Motor, Wi-fi	Rp.1.751.000 – Rp.2.000.000	Valid
1b	Lubuk Baja, 3 x 4, Termasuk Listrik, AC, WC Dalam, Kasur, Kloset Duduk, Kursi, Lemari, Meja, TV, Parkir Mobil, Parkir Motor, Wi-fi	Rp.1.751.000 – Rp.2.000.000	
2a	Nongsasari, 2 x 3, Air Saja, WC Dalam, Kasur, Kipas Angin, Lemari, TV, Dapur, Parkir Mobil, Parkir Motor, Wi-fi	Rp.500.000 – Rp.750.000	Valid
2b	Nongsasari, 2 x 3, Air Saja, WC Dalam, Kasur, Kipas Angin, Lemari, TV, Dapur, Parkir Mobil, Parkir Motor, Wi-fi	Rp.500.000 – Rp.750.000	
3a	Batam Kota, 4 x 4, Termasuk Listrik, AC, Air Panas, WC Dalam, Kasur, Kloset Duduk, Kursi, Lemari, Meja, TV, Dapur, WC Luar, Kulkas, Laundry, Mesin Cuci, Parkir Mobil, Parkir Motor, TV Luar, Ruang Santai, Wi-fi	Rp.2.251.000 – Rp.2.500.000	Valid
3b	Batam Kota, 4 x 4, Termasuk Listrik, AC, Air Panas, WC Dalam, Kasur, Kloset Duduk, Kursi, Lemari, Meja, TV, Dapur, WC Luar, Kulkas, Laundry, Mesin Cuci, Parkir Mobil, Parkir Motor, TV Luar, Ruang Santai, Wi-fi	Rp.2.251.000 – Rp.2.500.000	

Pada tabel di atas, kode a adalah untuk uji test di website sedangkan kode b untuk uji test di modelnya. Dari 3 uji test diatas semuanya terlihat memprediksikan harga yang sama dan bisa dinyatakan valid, yang berarti tidak ada kesalahan dalam melakukan pengkodean menggunakan flask.

IV. KESIMPULAN

Website prediksi harga kost ini bisa membantu masyarakat dalam memprediksi harga kost yang ingin mereka sewa maupun mereka sewakan. Dengan menggunakan algoritma dari supervised learning yaitu linear regression dan random forest maka pembuatan model prediksi untuk website bisa dikembangkan. Untuk algoritma linear regression mendapatkan nilai R2 Score sebesar 64% dan untuk algoritma random forest mendapatkan R2 Score sebesar 99% yang dimana berarti model dari algoritma random forest yang akan di gunakan untuk pengembangan website.

Harga dari sebuah kost sangat dipengaruhi dari fasilitas dalam kamar seperti adanya AC atau tidak atau sama sekali kamar kosong. Dan fasilitas dari luar kamar atau fasilitas bersama itu sangat di pengaruhi oleh adanya wc atau tidak yang berarti itu juga sangat berpengaruh ketika wc berada didalam kamar.

DAFTAR PUSTAKA

- [1] M. Sumarni and S. Wahyuni, "Determinan Keputusan Mahasiswa Ekonomi Syariah Memilih Rumah Kost (Studi Kasus Mahasiswa Prodi Ekonomi Syariah IAIN Langsa)," *J. Mhs. Akunt. SAMUDRA*, vol. 2, no. 3, pp. 230–240, 2021.
- [2] A. Suparwo, R. Roisah, A. Solihat, and Fitriyani, "Strategi Pemasaran Online Berbasis Aplikasi Pada Kost Di Wilayah Pasundan Bandung," *J. Pengabd. Kpd. Masy. Nusant.*, vol. 3, no. 2, pp. 1634–1642, 2023.
- [3] D. Handayani and H. Lubis, "Sistem Informasi Manajemen Aplikasi Rumah Kost dengan Menggunakan Rapid Application Development Berbasis Android dan SMS Gateway," *JSI (Jurnal Sist. Informasi) Univ. Suryadana*, vol. 8, no. 1, pp. 83–88, 2021, doi: <https://doi.org/10.35968/jsi.v8i1.610>.
- [4] N. S. A. Laily and Rapina, "Rancang Bangun Rumah Kost Berbasis Web Untuk Memudahkan Pencarian Kost Di Kota Batam," *Zo. Komput.*, vol. 11, no. 1, pp. 49–57, 2021.
- [5] D. Sunarsi, A. Barsah, H. Hastono, and I. R. Akbar, "Pengaruh Harga dan Fasilitas Terhadap Minat Sewa Kost yang Berdampak Pada Prestasi Belajar Mahasiswa di Yogyakarta," *J. Educ. Hum. Soc. Sci.*, vol. 3, no. 2, pp. 601–609, 2020, doi: [10.34007/jehss.v3i2.369](https://doi.org/10.34007/jehss.v3i2.369).
- [6] M. D. Ilhami and Y. Feri, "Pengaruh Harga dan Lokasi Terhadap Keputusan Sewa Kamar Kost (Studi Kasus Pada Mahasiswa yang Sewa Kamar Kost di Kawasan Adam Malik KM.9)," *J. Indones. Manag.*, vol. 3, no. 2, pp. 229–240, 2023.
- [7] L. Utari and A. Zulfikar, "Penerapan Convolutional Neural Networks Menggunakan Edge Detection Untuk Identifikasi Motif Jenis Batik," *TeknoIS J. Ilm. Teknol. Inf. dan Sains*, vol. 13, no. 1, pp. 110–123, 2023, doi: [10.36350/jbs.v13i1.184](https://doi.org/10.36350/jbs.v13i1.184).
- [8] L. U. Khasanah, "Kenali Pengertian Machine Learning & Contohnya, Pemula Catat!," *DQLab*, 2022. <https://dqlab.id/kenali-pengertian-machine-learning-and-contohnya-pemula-catat> (accessed Nov. 03, 2022).
- [9] F. R. Lumbanraja, R. A. Saputra, K. Muludi, A. Hijriani, and A. Junaidi, "Implementasi Support Vector Machine Dalam Memprediksi Harga Rumah Pada Perumahan Di Kota Bandar Lampung," *J. Pepadun*, vol. 2, no. 3, pp. 327–335, 2021, doi: [10.23960/pepadun.v2i3.90](https://doi.org/10.23960/pepadun.v2i3.90).
- [10] M. R. Fahlepi and A. Widjaja, "Penerapan Metode Multiple Linear Regression Untuk Prediksi Harga Sewa Kamar Kost," *J. Strateg.*, vol. 1, no. November, pp. 615–629, 2019.
- [11] K. M. B. Azhary, "Analisis Tren Kepemilikan Rumah di Kota Palembang dan Prediksi Harga Rumah memanfaatkan Machine Learning," *J. Perenc. Wil.*, vol. 8, no. 2, pp. 165–173, 2023, doi: [10.33772/jpw.v8i2.377](https://doi.org/10.33772/jpw.v8i2.377).
- [12] W. Ali and S. A. Aklani, "Analisis Algoritma Monte Carlo Untuk Memprediksi Keuntungan Pembangunan Apartemen Menggunakan SCRUM Framework," *J. Ilm. Betrik*, vol. 13, no. 03, pp. 287–294, 2022.
- [13] D. F. Ningtyas and N. Setiyawati, "Implementasi Flask Framework pada Pembangunan Aplikasi Purchasing Approval Request," *J. Janitra Inform. dan Sist. Inf.*, vol. 1, no. 1, pp. 19–34, 2021, doi: [10.25008/janitra.v1i1.120](https://doi.org/10.25008/janitra.v1i1.120).
- [14] U. Athiyah, A. Hananta, T. Maulidi, V. M. E. Putra, T. F. H. Purba, and A. A. W. Bakowatun, "Sistem Pendukung Keputusan Prediksi Harga Rumah Kost untuk Mahasiswa IT Telkom Purwokerto Menggunakan Metode Fuzzy Tsukamoto Berbasis Web," *J. Dinda Data Sci. Inf. Technol. Data Anal.*, vol. 1, no. 2, pp. 77–81, 2021, [Online]. Available: <http://journal.itelkompwt.ac.id/index.php/dinda>.
- [15] Trivusi, "Perbedaan MAE, MSE, RMSE, dan MAPE pada Data Science," *Trivusi*, 2023. https://www.trivusi.web.id/2023/03/perbedaan-mae-mse-rmse-dan-mape.html#google_vignette (accessed Oct. 29, 2023).
- [16] W. Supriyanti and D. A. Pertiwi, "Implementasi Scrum dalam Pengembangan Sistem Informasi Pengelolaan Nilai Siswa," *REMIK Ris. dan E-Jurnal Manaj. Inform. Komput.*, vol. 6, no. 3, pp. 547–560, 2022, [Online]. Available: <http://jurnal.polgan.ac.id/index.php/remik/article/view/11732>.