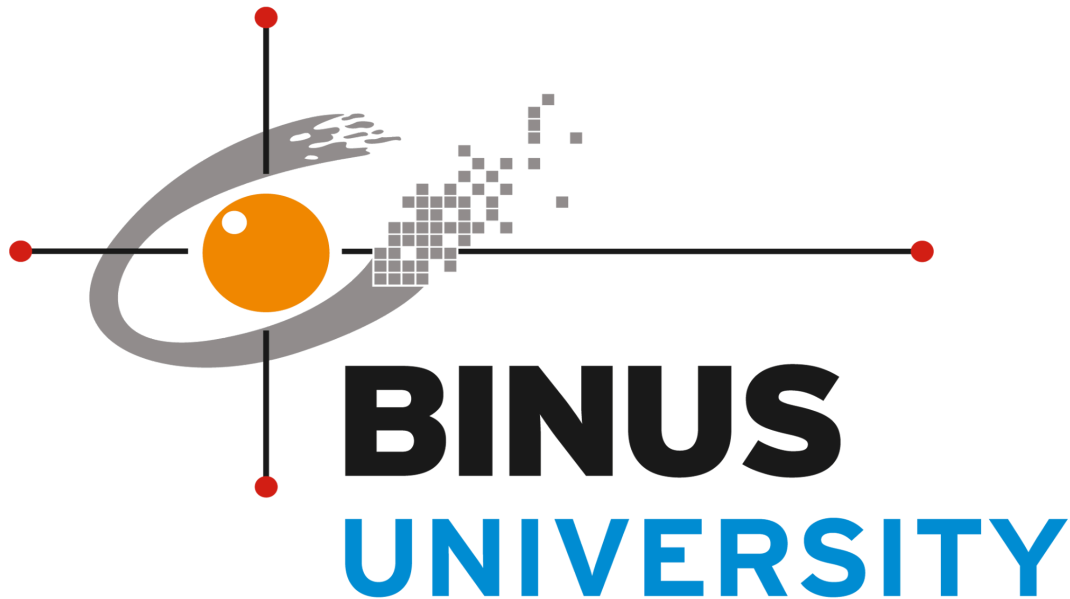


LAPORAN AKHIR ASSURANCE OF LEARNING
MATA KULIAH DATA ANALYTICS
BINUS UNIVERSITY GENAP 2024
KELAS LB01

“Job and Salaries in Data Science”

People
Innovation
Excellence



Disusun Oleh:

2602108262	Adytia Wiguna Gunawan	COMPUTER SCIENCE
2602141051	Ngui Su Ying	COMPUTER SCIENCE
2602136934	Hizkia Imanuel Septianto	COMPUTER SCIENCE

PROGRAM STUDI KEILMUAN COMPUTER SCIENCE
FAKULTAS SCHOOL OF COMPUTER SCIENCE

UNIVERSITAS BINA NUSANTARA

JAKARTA, 2024

DAFTAR ISI

DAFTAR ISI.....	3
BAB I.....	4
1.1 Latar Belakang.....	4
BAB II.....	7
2.1. Analisis Data.....	7
2.2. Ilmu Data.....	7
2.2.1. Profesi Ilmu Data.....	8
2.3. Himpunan Data.....	8
2.4. Atribut.....	9
2.5. Model Himpunan Data.....	10
BAB III.....	11
3.1. Metodologi.....	11
3.1.1. Regresi Linear.....	11
3.1.2. Decision Tree Regressor.....	12
3.1.3. Random Forest Classifier.....	12
BAB IV.....	14
4.1. Hasil Prediksi Model I.....	14
4.1 Hasil Prediksi Model II.....	15
4.2 Hasil Prediksi Model III.....	16
BAB V.....	18
KESIMPULAN.....	18
Daftar Pustaka.....	19

BAB I

PENDAHULUAN

1.1 Latar Belakang

Berkembangnya teknologi dan informatika juga memberikan perkembangan pada bidang industri, perkembangan industri dimulai dari industri 1.0 hingga saat ini yaitu industri 4.0 yang semakin canggih. Industri 4.0 merupakan revolusi industri yang menggunakan teknologi machine learning dan AI memungkinkan produksi yang cerdas, efisien, efektif, individual, dan disesuaikan dengan biaya yang terjangkau. Industri 4.0 membutuhkan bantuan komputer dan mesin yang pintar sehingga proses lebih cepat, sensor yang lebih kecil, penyimpanan data yang lebih murah, transmisi data yang lebih murah. Dalam industri 4.0, mesin dan produk saling berkomunikasi dan belajar dari satu sama lain sehingga mesin semakin pintar dalam menjalankan tugasnya. Perusahaan juga sudah banyak yang menerapkan industri 4.0 karena memberikan kemudahan bagi perusahaan dalam mengatasi tantangan dan masalah yang dimiliki. Industri 4.0 membutuhkan sumber daya manusia seperti data scientist dan ahli dalam tiap bidang yang terdapat pada industri 4.0 yaitu, IoT(Internet of Things), cybersecurity, robotics, Augmented Reality(AR), dan Artificial Intelligence(AI) agar penggunaan mesin jadi lebih efektif dan ahli dalam menganalisis data. Hal ini menjadi tantangan tersendiri dalam penerapan industri 4.0 karena sumber daya manusia yang terbatas tidak dapat memenuhi tingginya minat perusahaan.

Akhir-akhir ini, profesi *data scientist* cukup menjadi sorotan. *Data scientist* menjadi profesi terbaik di Amerika dengan gaji awal yang tertinggi dibandingkan posisi entry-level, dilansir dari laman glassdoor.com, seorang *data scientist* di Indonesia dapat memperoleh gaji Rp 7.000.000 hingga dua digit sebagai *entry-level*. Profesi ini terdiri dari tiga bidang ilmu yaitu bahasa pemrograman, matematika statistik, dan bisnis. *Data scientist* dibutuhkan setiap perusahaan maupun perusahaan kecil dan besar untuk mengolah data perusahaan yang dimiliki sehingga profesi ini memiliki permintaan yang sangat tinggi. *Data*

scientist berperan untuk mengolah data untuk mendukung penentuan keputusan bisnis dari perusahaan tersebut. Keputusan bisnis menjadi penentu untuk merencanakan strategi dan keuntungan perusahaan yang mendukung kesuksesan perusahaan.

Profesi *data scientist* sempat mengalami kendala karena semakin kesini semakin banyak juga *data scientist* yang keluar dari pekerjaannya. Dilansir dari laman [linkedin.com](https://www.linkedin.com) dan analyticsvidhya.com, hal ini dikarenakan kurangnya pengakuan dan apresiasi dalam bentuk gaji serta pembayaran gaji yang tidak jelas. Gaji yang tidak memuaskan dapat menyebabkan performa kerja yang tidak memuaskan serta lingkungan kerja menjadi tidak nyaman. Oleh karena itu diperlukan kejelasan pada gaji karyawan. *Data scientist* membutuhkan sistem pembayaran gaji yang adil dan transparan yang mencerminkan keterampilan, pengalaman, dan kinerja mereka. Masalah ini dapat diatasi dengan perencanaan kompensasi yang dilakukan oleh pekerja dan perusahaan.

Perencanaan kompensasi merupakan strategi merencanakan dan menentukan pembayaran kisaran gaji. Sebanyak 76% pekerja menerima pekerjaannya dengan mempertimbangkan kompensasi yang ditawarkan. Tujuan perencanaan kompensasi agar pekerja mendapat gaji yang adil berdasarkan performa kerja dan standar yang dimiliki pekerja. Strategi ini dapat mempertahankan pekerja dengan talenta terbaik serta mengurangi tingkat *turnover* yang menjadi tolok ukur kompetitif di pasar. Perencanaan kompensasi ditentukan oleh beberapa faktor yaitu pekerjaan, performa kerja, pengalaman kerja, dan analisa pasar. Perencanaan kompensasi pekerja dapat dianalisis sehingga kompensasi dapat ditentukan dari keputusan yang terdapat pada data.

Analisis data kompensasi pekerja dapat mengidentifikasi gaji pekerja berdasarkan performa kerja dan standar yang dimiliki pekerja secara transparan karena keputusan ditentukan berdasarkan data yang terdapat pada *database*. Analisis data mengurangi kesenjangan dan menyetarakan gaji sehingga pembagian gaji menjadi lebih adil. Analisis data kompensasi pekerja juga membantu perusahaan untuk mengevaluasi laba atas investasi (ROI) program kompensasi yang diterapkan kepada pekerja. Oleh karena itu, analisis data

kompensasi pekerja diperlukan dalam menentukan keputusan penawaran pekerjaan bagi *data scientist*.

Untuk mengetahui upah sesuai dengan pekerjaan-pekerjaan yang ada di bidang Data Science, data yang dilansir dari Kaggle.com berjudul "Jobs and Salaries in Data Science" dapat dijadikan bahan analisis data untuk mengatasi permasalahan kompensasi pekerja sehingga lebih transparan. Untuk memprediksi upah sesuai dengan pekerjaan-pekerjaan yang ada di bidang Data Science, perlu dibatasi dengan atribut-atribut yang saling berkorelasi. Atribut-atribut yang dimaksud berupa seperti tahun bekerja, bidang pekerjaan, nama profesi atau pekerjaan, mata uang sesuai tempat bekerja, jumlah upah yang diterima sesuai mata uang tempat bekerja, jumlah upah yang diterima dalam mata uang USD (\$), tempat tinggal pekerja (negara), pengalaman berapa lama bekerja, jenis pekerjaan (part-time, full-time, ataupun lainnya), dan cara bekerja (remote, in-person, ataupun lainnya).

BAB II

STUDI LITERATUR

2.1. Analisis Data

Analisis data merupakan ilmu untuk menganalisis data mentah sehingga dapat menghasilkan kesimpulan dari data yang dimiliki. Analisis data membutuhkan metode, teknik, perangkat lunak, dan alat untuk mengumpulkan data, memproses data mentah hingga menjadi informasi dari data tersebut. Analisis data mengandalkan data yang dimiliki maka data harus berkualitas. Data harus dapat diandalkan karena jika data yang dimiliki tidak akurat, tidak terstruktur, dan tidak lengkap maka hasil analisis data tidak terjamin keandalannya. Data yang diolah dalam kuantitas yang besar bersifat kompleks sehingga dibutuhkan pemahaman statistik yang kuat. Proses yang dilalui menghasilkan kesimpulan analisis data yang menjadi informasi berharga bagi suatu perusahaan.

2.2. Ilmu Data

Ilmu data merupakan teknik untuk menganalisis data yang melibatkan pembelajaran mesin untuk menghasilkan informasi dari data yang dimiliki. Ilmu data tersusun dari bidang matematika, statistika, Artificial Intelligence(AI), dan teknik komputer yang bertujuan untuk memproses data menjadi kesimpulan sebagai penentu keputusan suatu perusahaan. Menganalisis data membantu perusahaan menentukan keputusan dengan pemahaman yang baik terhadap perusahaan karena kesimpulan diambil dari hasil analisis data yang objektif. Hasil analisis data juga menjadi bahan pertimbangan perusahaan dalam mengidentifikasi pola dalam data. Analisis data tersusun atas data, variabel, hipotesis, statistik, metode analisis. Data mentah yang dimiliki tersusun atas variabel-variabel yang saling berkorelasi. Analisis data bertanggung jawab untuk membuat hipotesis dari variabel yang berkorelasi. Setelah itu, data akan diolah menjadi statistik yang menjadi alat untuk mengidentifikasi pola pada data yang dimiliki. Data harus dianalisis berdasarkan metode analisis yang sesuai dengan jenis dan tujuan analisis. Perusahaan dapat memaksimalkan keuntungan dengan menentukan strategi yang efisien

berdasarkan hasil analisis data. Oleh karena itu, ilmuwan data sangat diperlukan oleh perusahaan untuk membantu menentukan keputusan bisnis yang lebih optimal.

2.2.1. Profesi Ilmu Data

Tingginya permintaan perusahaan pada sumber daya ilmu data menjadikan ilmuwan data sebagai profesi yang menjanjikan. Ilmuwan data merupakan profesi yang umum di bidang ilmu data, biasanya ilmuwan data berperan sebagai pemimpin tim karena keahliannya. Ilmuwan data bertugas untuk mengidentifikasi bisnis, mengumpulkan data, menganalisa data, dan memvisualisasikan data. Ilmuwan data dan analisis data memiliki peran yang tidak jauh berbeda karena perusahaan bisa saja mempekerjakan seorang ilmuwan data dengan jobdesk seorang analisis data yaitu memvisualisasikan, mengubah, dan memanipulasi data sebagai bahan komunikasi bisnis. Bahan data yang dimiliki akan dibuat narasi untuk menggambarkan data yang akan dilakukan oleh pendongeng data. Dalam suatu bisnis, persyaratan model bisnis juga perlu dirancang dengan sistem database yang dipelihara dengan baik. Arsitek data bertugas untuk memelihara sistem dengan melacak data dan mengatur akses database. Profesi ilmu data dan tanggung jawab tiap profesi diilansir dari laman builtin.com.

2.3. Himpunan Data

Untuk melakukan penelitian tentang pekerjaan dan gaji dalam ilmu data, dimulai dengan mengumpulkan data yang relevan. Himpunan data untuk penelitian ini diilansir dari Kaggle.com dengan judul “Job and Salary in Data Science”. Sumber data telah melalui proses pengecekan dan diintegrasikan untuk membuat himpunan data yang berkualitas. Terdapat 9355 data yang terdapat pada himpunan data.

Data akan diolah untuk proses training data dan validasi data. 50% dari data untuk proses training data dalam machine learning. 30% data akan dijadikan bahan validasi pembelajaran model sehingga model dapat menggeneralisasi dan mengidentifikasi pola untuk memprediksi data baru. Proses training data dan validasi data memberikan gambaran algoritma pada machine learning sehingga machine learning memberikan output sesuai yang diharapkan ilmuwan data. 20% dari data akan digunakan untuk proses *testing* data.

2.4. Atribut

Himpunan data “Job and Salary in Data Science” terdiri dari 12 atribut yaitu *Work_year*, *Job_title*, *Job_category*, *Salary_currency*, *Salary*, *Salary_in_usd*, *Employee_residence*, *Experience_level*, *Employment_type*, *Work_setting*, *Company_location*, dan *Company_size*.

- **Work_year** : Tahun data tercatat
- **Job_title** : Detail bidang pekerjaan mencakup '*Data Scientist*', '*Data Engineer*', atau '*Data Analyst*'
- **Job_category** : Mengklasifikasi peran pekerjaan dalam kategori yang lebih luas mencakup '*Data Analysis*', '*Machine Learning*', '*Data Engineering*'
- **Salary_currency** : Mata uang gaji
- **Salary** : Gaji yang didapatkan.
- **Salary_in_usd** : Gaji yang didapatkan dengan mata uang USD(\$)
- **Employee_residence** : Tempat tinggal pekerja
- **Experience_level** : Pengalaman kerja dan tingkat profesionalitas karyawan mencakup '*Entry-level*', '*Mid-level*', '*Senior*', dan '*Executive*'.
- **Employment_type** : Menentukan jenis pekerjaan mencakup '*Full-time*', '*Part-time*', '*Contract*'
- **Work_setting** : Pengaturan atau lingkungan kerja, seperti '*Remote*', '*In-person*', atau '*Hybrid*'
- **Company_location** : Lokasi negara perusahaan
- **Company_size** : Ukuran perusahaan yaitu small (S), medium (M), and large (L)

2.5. Model Himpunan Data

- **Work_year** : Tahun data tercatat untuk mengetahui tren gaji sesuai tahun data tercatat.
- **Job_title** : Detail bidang pekerjaan untuk mengetahui distribusi gaji sesuai bidangnya.
- **Job_category** : Mengklasifikasi peran pekerjaan dalam kategori untuk mempermudah analisa data.
- **Salary_currency** : Mata uang gaji untuk membantu membandingkan gaji secara global.
- **Salary** : Gaji yang didapatkan.
- **Salary_in_usd** : Gaji yang didapatkan dengan mata uang USD agar gaji yang didapatkan senilai sehingga mempermudah untuk dibandingkan secara global.
- **Employee_residence** : Tempat tinggal pekerja untuk mengetahui perbedaan gaji dan biaya hidup secara geografis.
- **Experience_level** : Pengalaman kerja dan tingkat profesionalitas karyawan untuk mengetahui bagaimana pengalaman mempengaruhi gaji.
- **Employment_type** : Menentukan jenis pekerjaan untuk menganalisis bagaimana pengaturan pekerjaan yang berbeda dapat memengaruhi struktur gaji.
- **Work_setting** : Pengaturan atau lingkungan kerja untuk memengaruhi tingkat gaji di industri data.
- **Company_location** : Lokasi negara perusahaan untuk menganalisis bagaimana negara perusahaan dapat memengaruhi struktur gaji.
- **Company_size** : Ukuran perusahaan untuk menganalisis bagaimana ukuran perusahaan memengaruhi gaji.

BAB III

PEMBAHASAN

3.1. Metodologi

Metode-metode yang digunakan untuk model yang dibuat pada penelitian ini adalah Regresi Linear, *Decision Tree Regressor*, dan *Random Forest Classifier*.

3.1.1. Regresi Linear

Regresi Linear merupakan metode yang digunakan untuk memodelkan hubungan linier antara satu atau beberapa variabel independen (variabel yang mempengaruhi) dengan variabel dependen (variabel yang ingin diprediksi). Dalam konteks penelitian ini, regresi linear memodelkan hubungan antara faktor-faktor seperti pengalaman, pendidikan, dan lokasi geografis dengan gaji dalam industri *Data Science*.

Proses regresi linear dimulai dengan mengestimasi koefisien regresi yang merupakan garis lurus terbaik yang sesuai dengan data. Estimasi ini dilakukan dengan meminimalkan jumlah kuadrat dari selisih antara nilai sebenarnya dan nilai prediksi (yang dihitung menggunakan model regresi) dari variabel dependen. Setelah model regresi dibangun, koefisien regresi yang dihasilkan dapat diinterpretasikan untuk memahami seberapa besar pengaruh setiap faktor terhadap gaji.

Dalam penelitian ini, regresi linear digunakan sebagai metode analisis utama untuk memahami hubungan antara pekerjaan dan gaji dalam industri *Data Science*. Regresi linear dipilih karena memungkinkan untuk membuat model dari hubungan linier antara faktor-faktor seperti pengalaman, pendidikan, lokasi geografis, dan faktor-faktor lain berdasarkan dataset yang ada dengan gaji, yang merupakan fokus utama penelitian ini. Selain itu, regresi linear memberikan kemudahan untuk interpretasi koefisien regresi, memungkinkan untuk memahami seberapa besar pengaruh setiap faktor terhadap gaji dalam industri ini.

3.1.2. *Decision Tree Regressor*

Decision Tree Regressor merupakan metode *machine learning* yang bertujuan untuk memprediksi nilai kontinu berdasarkan serangkaian aturan keputusan yang dibangun dalam bentuk struktur pohon. Tahap awal dari metode ini dimulai dengan membagi dataset menjadi subset-subset yang semakin kecil berdasarkan nilai dari fitur-fitur yang ada. Pemilihan fitur yang paling baik dilakukan dengan memilih kriteria yang paling efektif dalam memisahkan nilai target di dalam setiap subset, seperti *Mean Squared Error* (MSE) atau *Mean Absolute Error* (MAE). Proses ini akan terus berlanjut untuk membangun struktur *decision tree*, di mana setiap node merupakan decision berdasarkan fitur-fitur tertentu dan setiap cabang mewakili kemungkinan nilai dari fitur tersebut. Kemudian, data baru dapat diprediksi dengan di-*run* melalui *decision tree* yang sudah dibuat. Prediksi akhir diberikan berdasarkan nilai rata-rata atau median dari nilai target di node-terminal yang sesuai.

Model prediksi yang dibuat menggunakan metode *Decision Tree Regressor* karena beberapa hal. Pertama, hubungan antara atribut-atribut seperti pengalaman, lokasi geografis, dan lokasi perusahaan dengan gaji tidak selalu memiliki pola yang linier, maka dari itu *Decision Tree Regressor* digunakan karena mampu menangkap pola non-linier yang kompleks pada dataset tersebut. Kedua, walau tidak seinterpretatif seperti metode *Regresi Linear*, metode *Decision Tree Regressor* digunakan untuk memperjelas hasil prediksi dari metode *Regresi Linear*. Tidak hanya itu, metode ini juga dapat menangani data-data outlier dan hanya memerlukan sedikit *preprocessing* data.

3.1.3. *Random Forest Classifier*

Algoritma pembelajaran mesin yang dikembangkan oleh Leo Breiman dan Adele Cutler menggabungkan output dari beberapa *decision tree* untuk memperoleh hasil. Menggabungkan output dari beberapa *decision tree* meningkatkan prediksi yang lebih akurat. Algoritma ini memiliki kemampuan beradaptasi yang baik sehingga dapat mengatasi masalah klasifikasi dan regresi dengan baik.

Algoritma ini menggabungkan beberapa decision tree sehingga menggunakan teknik ensemble. Ensemble menggunakan dua jenis metode yaitu bagging(mengantongi) dan meningkatkan. Metode Sampel data akan di training sehingga menghasilkan training data yang berbeda dan output berdasarkan dari hasil voting. Setelah itu, data akan digabung sehingga menghasilkan model dengan akurasi yang tinggi.

Algoritma ini dimulai dengan mengumpulkan data yang kompleks, memitigasi overfitting, dan data akan diolah sehingga menjadi alat untuk memprediksi dalam machine learning. Metode ini dipilih menyesuaikan dengan dataset yang dimiliki. Pada klasifikasi, data yang dikumpulkan bersifat variabel kategori yaitu variabel yang diberikan label dan dapat masuk dikategorikan. Sedangkan regresi, data yang dikumpulkan bersifat variabel kontinu dimana suatu variabel yang nilainya diperoleh melalui pengukuran yang tak terhitung banyaknya pada. Atribut pada himpunan data yaitu *salary*, *work_year*, *salary_in_usd*, dan *experience_level* dikategorikan sebagai variabel kontinu dalam regresi. Atribut *salary* dan *salary_in_usd* dikategorikan sebagai variabel kontinu karena *salary* merupakan variabel dependen yang ingin diprediksi berdasarkan variabel independen lain yang terdapat pada himpunan data. Atribut *work_year* merupakan tahun data tercatat untuk menganalisa tren gaji yang dikategorikan sebagai variabel independen yang kontinu. Atribut *experience_level* merupakan pengalaman kerja dan tingkat profesionalitas karyawan sehingga dapat dikategorikan sebagai variabel kontinu.

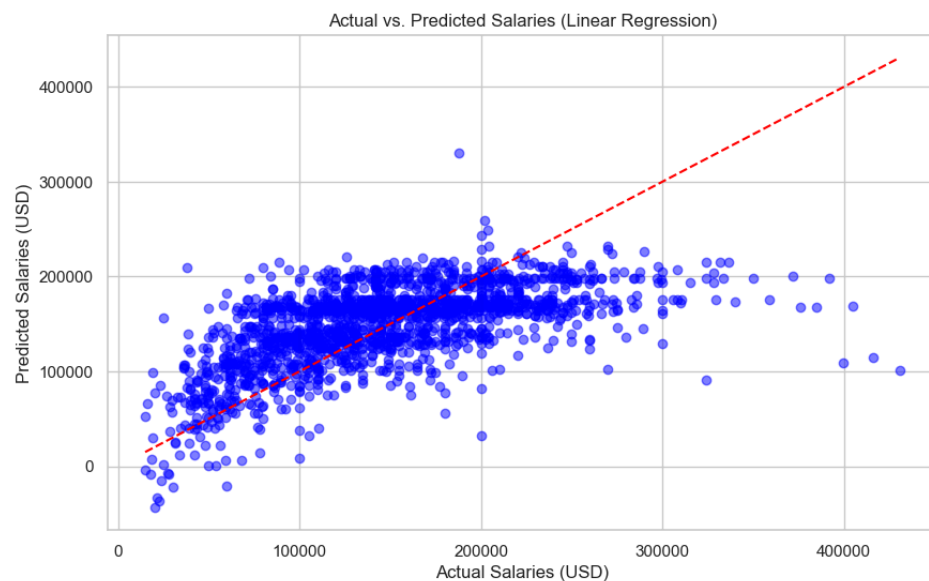
BAB IV

HASIL PEMBAHASAN

Pada penelitian ini, terdapat 3 model yang dibuat. Model yang pertama menggunakan metode *Regresi Linear*, model yang kedua menggunakan metode *Random Forest Classifier*, dan model yang ketiga menggunakan metode *Regresi Linear* dan *Decision Tree Regressor*. Berikut merupakan hasil yang diberikan dari masing-masing model prediksi yang dibuat di penelitian ini.

4.1. Hasil Prediksi Model I

Pada model I, metode yang digunakan adalah metode Regresi Linear. Dengan menggunakan bahasa pemrograman Python, model I menghasilkan hasil seperti berikut.



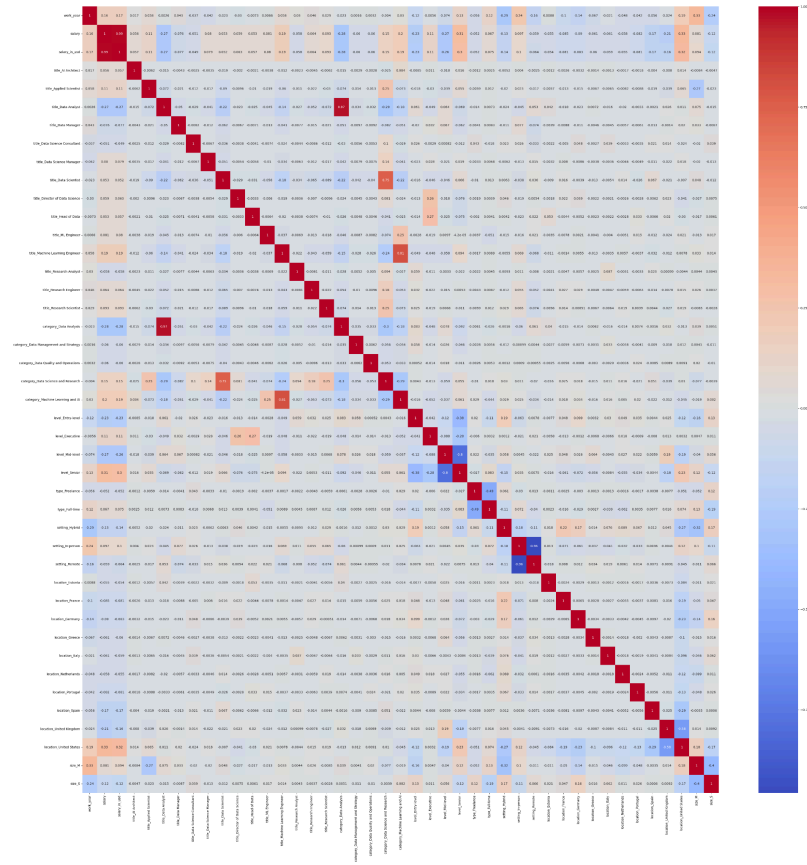
Hasil dari model I ini memiliki akurasi sebesar 35%. Seperti yang dapat dilihat pada gambar berikut.

```
regression_model.score(x_test,y_test)

0.35089786640706633
```

4.1 Hasil Prediksi Model II

Pada model II, metode yang digunakan adalah metode *Random Forest Classifier*. Dengan menggunakan bahasa pemrograman Python, model II menghasilkan hasil seperti berikut.



Hasil dari model II ini memiliki akurasi sebesar 22%. Seperti yang dapat dilihat pada gambar berikut.

```
forest.score(test_x_encoded, test_y)
```

```
0.2205628045904732
```

4.2 Hasil Prediksi Model III

Pada model III, metode yang digunakan adalah metode Regresi Linear dan *Decision Tree Regressor*. Dengan menggunakan bahasa pemrograman Python, model III menghasilkan hasil seperti berikut.

Hasil dari metode Regresi Linear.

```
from sklearn.metrics import mean_squared_error, mean_absolute_error
y_pred = linear_model.predict(X)
error = np.sqrt(mean_squared_error(Y, y_pred))
print("${:,.02f}".format(error))
```

```
$50,896.92
```

Hasil dari metode *Decision Tree Regressor*.

```
y_pred = dec_tree_reg.predict(X)
y_pred = dec_tree_reg.predict(X)
error = np.sqrt(mean_squared_error(Y, y_pred))
print("${:,.02f}".format(error))
```

```
$44,755.94
```

Seperti yang dapat dilihat pada gambar-gambar tersebut, *error* yang dihasilkan dari metode *Decision Tree Regressor* memiliki nilai yang lebih kecil daripada hasil *error* yang dihasilkan metode Regresi Linear. Dalam hal

ini, metode *Decision Tree Regressor* lebih efektif dan akurat dibandingkan dengan metode Regresi Linear.

BAB V

KESIMPULAN

Analisis data merupakan ilmu untuk menganalisis data mentah sehingga dapat menghasilkan kesimpulan dari data yang dimiliki. Pada kasus ini, penelitian yang dilakukan adalah untuk memprediksi gaji yang mungkin didapatkan, di industri *Data Science*. Dataset yang digunakan memiliki beberapa atribut, atribut-atribut tersebut merupakan faktor-faktor yang mungkin mempengaruhi jumlah gaji yang didapat. Pada penelitian ini, terdapat 3 model yang dibuat dengan 3 metode yang berbeda. Model I menggunakan metode Regresi Linear, yang dimana hasil prediksi dari model yang dibuat memiliki tingkat akurasi sebesar 35%. Sementara itu, Model II menggunakan metode *Random Forest Classifier*, dan tingkat akurasi prediksi yang dihasilkan sebesar 22%. Terakhir, Model III dibentuk menggunakan 2 metode, yaitu metode *Regresi Linear* dan metode *Decision Tree Regressor*. Pada Model III ini, hasil prediksi yang didapatkan memiliki *margin error* sebesar \$50,896.92 untuk hasil prediksi yang menggunakan metode Regresi Linear, dan sebesar \$44,755.94 untuk hasil prediksi yang menggunakan metode *Decision Tree Regressor*. Dari penerap ketiga model ini, dapat disimpulkan bahwa model yang paling bagus untuk digunakan diantara ketiga model yang telah dibuat adalah Model III, karena model ini memiliki tingkat akurasi yang lebih tinggi jika dibandingkan dengan yang lain.

Daftar Pustaka

Artikel Jurnal

- Alsudais, Abdulkareem, et al. "Comparison of Job Titles for Specific Terms: Investigating "Data Science."" 2022. https://www.researchgate.net/publication/365609334_Comparison_of_Job_Titles_for_Specific_Terms_Investigating_Data_Science. Diakses pada tanggal 3 April 2024.
- Quan, Tee Zhen, and Mafas Raheem. "Human Resource Analytics on Data Science Employment Based on Specialized Skill Sets with Salary Prediction." *International Journal of Data Science*, vol. 4, no. 1, 2023. <https://ijods.org/index.php/ds/article/view/64>. Diakses pada tanggal 3 April 2024.
- Vaidya, Saurabh, et al. "Industry 4.0 – A Glimpse." *2nd International Conference on Materials Manufacturing and Design Engineering*, 2018. https://www.sciencedirect.com/science/article/pii/S2351978918300672?ref=pdf_download&fr=RR-2&rr=86ea026e6ec968de. Diakses pada tanggal 3 April 2024.

Website

- "Apa itu Regresi Linier? - Penjelasan tentang Model Regresi Linear - AWS." *Amazon AWS*, <https://aws.amazon.com/id/what-is/linear-regression/>. Accessed 3 April 2024.
- "Compensation Planning: Comprehensive Guide+ 6 Step Process." *CHRMP*, <https://www.chrmp.com/compensation-planning/>. Diakses pada tanggal 3 April 2024.
- guide, step. "Compensation Plan: Benefits & actionable steps." *Leapsome*, <https://www.leapsome.com/blog/compensation-planning#why-compensation-planning-is-important>. Diakses pada tanggal 3 April 2024.
- Jain, Sneha. "5 Key Reasons Why Data Scientists Are Quitting their Jobs." *Analytics Vidhya*, 2020, <https://www.analyticsvidhya.com/blog/2019/12/5-key-reasons-data-scientists-quit-jobs/>. Diakses pada tanggal 3 April 2024.
- Rusinowitz, Sharon. "Why You Need to Use Data to Inform Your Compensation Strategy." *ChartHop*, 5 July 2023,

<https://www.charthop.com/resources/blog/comp-management/data-compensation-strategy/>. Diakses pada tanggal 3 April 2024.

“Salary: Data Scientist in Jakarta (Indonesia) 2024.” *Glassdoor*, https://www.glassdoor.com/Salaries/jakarta-data-scientist-salary-SRCH_IL.0,7_IC2709872_KO8,22.htm. Diakses pada tanggal 3 April 2024.

Sigma Team. “Hate Data Science? 3 Exit Opportunities for Data Scientists | Sigma Computing.” *Sigma Computing*, 18 June 2020, <https://www.sigmacomputing.com/blog/why-data-scientists-are-eyeing-the-exit>. Diakses pada tanggal 3 April 2024.

“Top Reasons Data Scientists Quit Their Jobs.” *LinkedIn*, 8 December 2023, <https://www.linkedin.com/advice/0/what-top-reasons-data-scientists-leave-jobs-skills-data-science-a689f>. Diakses pada tanggal 3 April 2024.