



<http://jurnal.universitaspurabangsa.ac.id/index.php/ijasta>

e-ISSN: 2829-4858

ARTICLE INFORMATION

Received March 18th 2023

Accepted March 30th 2023

Published April 1st 2023

**PERBANDINGAN AKURASI PREDIKSI METODE
REGRESI LINEAR OLS DENGAN JARINGAN SYARAF
TIRUAN : STUDI KASUS DATASET STARTUP**

Wahyuni Windasari¹, Wasiman², M. Annas Alfiyanto³

^{1,2,3}Universitas Putra Bangsa

email: wahyuwindasari@gmail.com

ABSTRAK

Perkembangan digital membawa pengaruh pada tumbuhnya perusahaan StartUp. Penelitian terkait dengan faktor yang mempengaruhi profit perusahaan StartUp menjadi bahan kajian yang menarik untuk dikaji lebih lanjut. Pada penelitian ini akan dilakukan analisis terkait faktor-faktor yang mempengaruhi profit perusahaan StartUp menggunakan metode prediksi linear berganda dan Jaringan Syaraf Tiruan (JST). Dengan melakukan analisis data pada dataset Kaggle dari 1000 perusahaan StartUp menggunakan RStudio diperoleh hasil bahwa performa model linear berganda lebih akurat dibandingkan dengan performa model JST. Hal ini terlihat dari nilai SAE model linear berganda lebih kecil dibandingkan SAE model JST.

Kata Kunci: StartUp; Regresi Berganda; Jaringan Syaraf Tiruan ; Neural Network

ABSTRACT

Digital developments have an impact on the growth of StartUp companies. Research related to the factors that affect the profit of the StartUp company is an interesting study material for further study. In this study, an analysis will be carried out related to StartUp company profits using multiple linear prediction methods and Artificial Neural Networks (ANN). By analyzing data on the Kaggle dataset from 1000 StartUp companies using RStudio, the results show that the performance of the multiple linear model is more accurate than the performance of the ANN model. This can be seen from the SAE value of the multiple linear model which is smaller than the SAE of the ANN model.

Keywords: StartUp; Multiple Linear Regression ; Neural Network

PENDAHULUAN

Pandemi Covid-19 membawa dampak di berbagai sektor kehidupan. Beberapa kebijakan seperti pembatasan aktivitas masyarakat seperti *physical distancing* hingga *lockdown* diberlakukan untuk mengurangi penyebaran virus Covid-19. Kebijakan-kebijakan tersebut memberikan dampak pada semua sektor, tak terkecuali dengan gaya hidup masyarakat yang semula hampir semua kegiatan dilakukan

secara konvensional beralih menjadi online. Perubahan ke arah digital memberikan pengaruh yang luar biasa pada perkembangan perusahaan Startup. Perusahaan StartUp merupakan perusahaan rintisan yang mampu berinovasi dengan teknologi digital sehingga keberadannya dinilai mampu menjawab kebutuhan masyarakat pada masa Pandemi Covid-19. Berdasarkan data Startup Ranking (StartupRanking, 2022), negara yang menduduki peringkat pertama dengan jumlah StartUp terbanyak di dunia tahun 2022 diduduki oleh Amerika Serikat. Tercatat setidaknya terdapat 77.082 perusahaan rintisan di negara tersebut. Beberapa produk StartUp terkenal dari Negara Amerika Serikat antara lain Coursera, Teespring, dan Giphy.

Sebagai perusahaan yang masuk ke dalam kategori usaha yang baru dirintis, waktu operasi perusahaan StartUp terhitung masih belum lama. Selain itu, perusahaan-perusahaan ini juga sebagian besar masih tergolong pada fase pengembangan dan penelitian untuk menemukan pasar yang tepat. Oleh karena itu, melakukan analisis terkait faktor yang mempengaruhi profit perusahaan StartUp menjadi salah satu hal yang penting untuk diperhatikan. Ada beberapa metode yang dapat digunakan untuk menjelaskan hubungan variabel terikat dari faktor-faktor lain yang mempengaruhi. Dalam artikel ini, penulis akan menerapkan metode regresi linear berganda dan Jaringan Syaraf Tiruan (JST) untuk menjelaskan hubungan variabel terikat terhadap variabel bebasnya. Metode analisis yang baik adalah metode analisis yang memberikan nilai kesalahan yang lebih kecil jika dibandingkan dengan metode analisis lainnya. Terdapat beberapa kajian terkait perbandingan kinerja metode analisis regresi berganda dan JST. Hasil penelitian Amrin(2016), menunjukkan bahwa kinerja metode regresi linear berganda lebih baik dibandingkan dengan Jaringan Syaraf Tiruan Backpropagation dalam memprediksi nilai inflasi di Indonesia. Hal ini berbanding terbalik dengan hasil penelitian dari Nafi'iyah (2016) yang menunjukkan bahwa kinerja metode JST Backpropagation lebih akurat dibandingkan dengan metode regresi linear dalam memprediksi harga emas.

Berdasarkan atas perbedaan hasil penelitian tersebut, maka pada penelitian ini akan dikaji terkait perbandingan akurasi metode regresi berganda dan JST dalam memprediksi besaran profit perusahaan StartUp di Amerika Serikat.

METODE

Penelitian ini menggunakan data sekunder yang bersumber dari Kaggle DataSet (Rupak Roy, n.d.), yang berisi data dari 1000 perusahaan Start Up di Amerika Serikat. Dari total keseluruhan data dibagi menjadi 2 bagian yaitu sebanyak 995 data sebagai data training dan 5 data sebagai data testing. Adapun perangkat lunak yang digunakan untuk olah data pada penelitian ini menggunakan software R.

Selanjutnya, untuk langkah-langkah pada penelitian ini adalah sebagai berikut :

1. Analisis regresi linear berganda

Regresi linear berganda merupakan analisis statistik yang memodelkan hubungan diantara variabel terikat dan variabel bebas menurut bentuk persamaan linear. Adapun bentuk persamaan linear pada penelitian ini adalah :

$$Y = a + bX_1 + cX_2 + dX_3 + e \quad [1]$$

Dimana ;

Y : Profit

X_1 : Biaya R&D

X_2 : Biaya Administrasi

X_3 : Biaya Marketing.

2. Analisis Jaringan Syaraf Tiruan

Jaringan Syaraf Tiruan (JST) memiliki struktur model yang sama dengan struktur model pada regresi linear berganda yaitu menggunakan satu variabel terikat dan beberapa variabel bebas. Hanya saja

pada metode JST tidak memerlukan penjelasan secara teoritis terkait dengan pengaruh antar variabel (Wijayanti, 2019). JST memetakan data input dan output menggunakan konsep belajar. Dengan menggunakan input data masa lalu, JST melakukan pembelajaran sekaligus melakukan generalisasi terhadap karakteristik tingkah laku objek untuk menentukan output dari model tersebut (Nurmahaludin, 2014).

3. Pemilihan model terbaik.

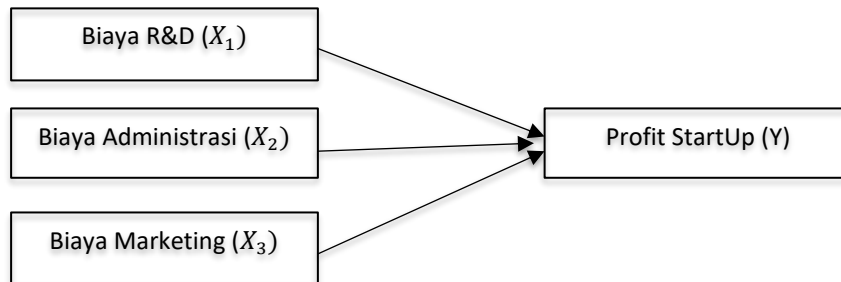
Model yang baik adalah model yang mampu menggambarkan data asli. Hal ini dapat dilihat dari jumlahan dari nilai absolute dari error model (SAE) yang terkecil.

$$SAE = \sum_{i=1}^n |X_i - X_t| \quad [2]$$

HASIL DAN PEMBAHASAN

1. Analisis Regresi Berganda

Variabel yang digunakan pada penelitian ini meliputi variabel profit sebagai variabel terikat, serta biaya R&D, biaya administrasi, serta biaya marketing sebagai variabel bebas. Adapun model penelitian yang dibangun diberikan pada Gambar 1. berikut ini ;



Gambar 1. Model Penelitian

Untuk pengolahan data regresi linear berganda menggunakan fungsi lm pada RStudio, berikut diberikan output analisis regresi berganda :

```

R 4.1.3 · ~/
> regresi <-lm(Profit ~RD + Administrasi +Marketing, data=data_Training)
> summary(regresi)

Call:
lm(formula = Profit ~ RD + Administrasi + Marketing, data = data_Training)

Residuals:
    Min       1Q   Median       3Q      Max
-353619 -153671 -44045   53620  732746

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.915e+05  1.398e+05   2.801  0.00519 **
RD          -1.455e+00  3.475e-01  -4.188  3.06e-05 ***
Administrasi -1.568e-01  1.245e+00  -0.126  0.89979
Marketing    -1.080e-01  1.276e-01  -0.846  0.39765
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 235700 on 991 degrees of freedom
Multiple R-squared:  0.09914, Adjusted R-squared:  0.09641
F-statistic: 36.35 on 3 and 991 DF, p-value: < 2.2e-16
  
```

Gambar 2. Output Fungsi lm pada RStudio

Dari output RStudio dapat diperoleh persamaan regresi berganda :

$$Profit = 391519,8 - 1,455RD - 0,157Administrasi - 0,108Marketing \quad [3]$$

Selanjutnya, dari model persamaan [2] yang sudah terbentuk diterapkan pada data testing. Adapun hasilnya diberikan pada tabel berikut :

PERBANDINGAN HASIL PREDIKSI METODE REGRESI LINEAR OLS DENGAN JARINGAN SYARAF TIRUAN:
STUDI KASUS DATASET STARTUP

Tabel 1.
Prediksi Profit StartUp dengan Regresi Berganda

No	Aktual	Prediksi	Error
996	95279,96	275458,35	-180178,38
997	164336,6	139103,15	25233,45
998	413956,5	183126,42	230830,06
999	333962,2	123723,11	210239,08
1000	476485,4	82562,541	393922,89

2. Analisis Jaringan Syaraf Tiruan

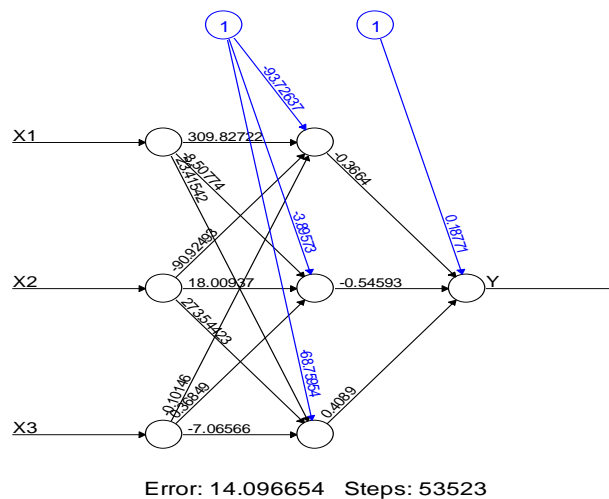
Pada tahap awal dilakukan proses pelatihan JST menggunakan 995 data dengan fungsi neuralnet pada RStudio. Analisis JST ini menggunakan arsitektur jaringan yang terdiri dari :

- a. Lapisan input dengan 3 neuron, yaitu biaya R&D, biaya administrasi, serta biaya marketing,
- b. Lapisan output dengan 1 neuron, yaitu Profit StartUp
- c. Satu lapisan tersembunyi dengan jumlah neuron yang berbeda-beda.

Penentuan arsitektur terbaik berdasarkan dari nilai error terkecil yang diperoleh dari arsitektur JST yang terbentuk yang dilakukan secara trial-error. Berikut diberikan hasil pelatihan jaringan dengan beberapa variasi jumlah neuron pada lapisan tersembunyi.

Tabel 2.
Perbandingan Pelatihan Arsitektur JST

Arsitektur Jaringan	Error
3-1-1	15,457
3-2-1	14,416
3-3-1	14,097
3-4-1	14,599
3-5-1	14,532



Gambar 3. Arsitektur JST 3-3-1

Dari hasil trial-error konstruksi arsitektur JST, diperoleh arsitektur 3-3-1 memberikan nilai kesalahan paling kecil. Selanjutnya, dengan menggunakan arsitektur terpilih dilakukan prediksi untuk data testing. Adapun hasil akhir dari nilai prediksi diberikan pada output R berikut :

```
> nnmodel <- neuralnet(Y ~X1+X2+X3, data=data_Training, hidden=c(3),linear.output=T)
> result_nn <-compute(nnmodel, data_Testing[,7:9])
> result_nn [["net.result"]]
      [,1]
[1,] 0.5051490
[2,] 0.2200106
[3,] 0.2475721
[4,] 0.2552807
[5,] 0.2551102
```

Gambar 4. Hasil Output R untuk Prediksi Data Testing

Untuk mendapatkan prediksi nilai yang sesungguhnya maka perlu di denormalisasikan kembali. Adapun hasil prediksi dan error dari analisis JST diberikan pada Tabel berikut.

Tabel 3.
Prediksi Profit StartUp dengan Analisis JST

No	Prediksi JST	Prediksi	Error
996	0,505	510293,664	-415013,701
997	0,220	157329,267	7007,339
998	0,248	191446,833	222509,647
999	0,255	200989,081	132973,109
1000	0,255	200778,024	275707,406

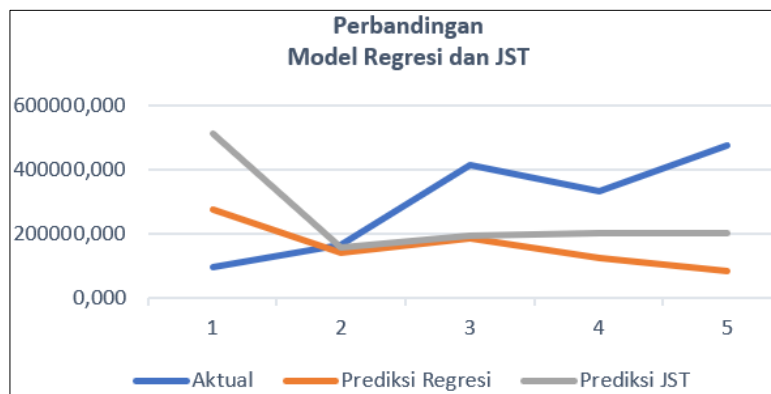
3. Pemilihan model terbaik

Pemilihan model terbaik didasarkan pada hasil perbandingan kinerja model. Model yang terbaik adalah model yang memberikan nilai kesalahan prediksi yang terkecil.

Tabel 4.
Komparasi Kinerja Model

Model	SAE
Regresi Berganda	1040404
JST	1053211

Dari hasil perbandingan pada Tabel 4 diperoleh bahwa nilai SAE model regresi berganda lebih kecil dibandingkan dengan SAE dari analisis JST. Hal ini menunjukkan bahwa model regresi berganda memiliki tingkat akurasi yang lebih baik dibandingkan analisis JST dalam hal memprediksi nilai profit dari perusahaan StartUp. Hal ini juga terlihat dari perbandingan performa kinerja regresi dan JST pada Gambar 4 berikut. Grafik dari prediksi regresi memiliki nilai fluktuasi yang lebih cenderung dapat mengikuti pergerakan dari grafik data asli dibandingkan dengan grafik hasil prediksi dari metode JST.



Gambar 4. Grafik Performa Model Regresi dan JST

SIMPULAN

Berdasarkan hasil analisis dapat disimpulkan :

1. Performa model regresi dalam memprediksi profit perusahaan StartUp memberikan nilai akurasi SAE sebesar 1040404.
2. Performa model JST dalam memprediksi profit perusahaan StartUp memberikan nilai akurasi model SAE sebesar 1053211.
3. Performa model linear berganda lebih baik dibandingkan dengan model JST dalam memprediksi profit StartUp.

REFERENSI

- Amrin, A. (2016). Analisa Komparasi Neural Network Backpropagation Dan Multiple Linear Regression Untuk Peramalan Tingkat Inflasi. *Jurnal Teknik Komputer AMIK BSI*, 2(2), 2442–2436. <https://doi.org/https://doi.org/10.31294/jtk.v2i2.1591>
- Nafi'iyah, N. (2016). Perbandingan Regresi Linear , Backpropagation Dan Fuzzy Mamdani Dalam Prediksi Harga Emas. *Seminar Nasional Inovasi Dan Aplikasi Teknologi Di Industri*, 291–296.
- Nurmahaludin. (2014). Analisis Perbandingan Metode Jaringan Saraf Tiruan dan Regresi Linear Berganda pada Prakiraan Cuaca. *Jurnal INTEKNA*, 2, 102–209. <https://ejurnal.poliban.ac.id/index.php/intekna/article/view/389>
- Rupak Roy. (n.d.). *1000_companies_profit*. Retrieved March 14, 2023, from <https://www.kaggle.com/datasets/rupakroy/1000-companies-profit>
- StartupRanking. (2022). *No Title*. <https://www.startupranking.com/countries>
- Wijayanti, A. E. S. A. (2019). Perbandingan Analisis Regresi Linear Berganda dan Jaringan Syaraf Tiruan Untuk Prediksi Nilai Tukar Rupiah. In *UIN Sunan Kalijaga* (Vol. 8, Issue 5). UIN Sunan Kalijaga.