

Pendekatan Data Science Dengan Algoritma Neural Network Dalam Penyusunan Strategi Pemberantasan Buta Aksara Dan Berhitung Di Negara-Negara Asean

Nur Sucahyo ^{*)1)}, Ike Kurniati²⁾, Syarif Hidayatullah³⁾

^{1)2) 3)}Program Studi Sistem Informasi, Institut Teknologi dan Bisnis Swadharma

^{*)}Correspondence Author: nursucahyo@swadharma.ac.id, DKI Jakarta, Indonesia

DOI: <https://doi.org/10.37012/jtik.v6i2.293>

Abstrak

Pendidikan menjadi suatu hal yang penting bagi setiap individu dan menjadi penentu bagi perkembangan negara. Berdasarkan data ASEAN Yearbook 2015, Singapura merupakan negara dengan nilai indeks pendidikan tertinggi dengan skor 0,768. Singapura tidak hanya memiliki sistem pendidikan berkualitas terbaik di ASEAN, tapi juga dunia. Dengan nilai Indeks Pendidikan alias EDI sebesar 0,692, Brunei Darussalam menempati posisi 30 di dunia dan nomor dua di Asia Tenggara. Malaysia menempati posisi 62 dalam daftar pendidikan terbaik didunia dan ketiga di ASEAN dengan skor EDI 0,671. Untuk menurunkan angka buta huruf dibutuhkan poin-poin strategis untuk menaikkan angka melek huruf atau literacy rate di negara-negara ASEAN. Metode penelitian yang dilakukan adalah dengan cara memanfaatkan data yang disediakan ASEAN DSE sebagai data primer serta data world bank dan UNESCO sebagai data sekunder. Pengolahan data menggunakan IBM SPSS Modeler. Algoritma Neural Network digunakan untuk mengetahui signifikan faktor yang mempengaruhi literacy rate di ASEAN. Dari penelitian didapatkan hasil berupa faktor signifikan yang mempengaruhi literacy rate di ASEAN yaitu : Angka anggaran pemerintah pada bidang pendidikan, rasio jumlah guru dengan jumlah murid, rasio jumlah murid perempuan dibandingkan dengan jumlah murid laki-laki, angka populasi yang bersekolah disekolah dasar dan menengah berdasarkan tempat tinggal, dan populasi yang bersekolah disekolah dasar dan menengah berdasarkan tingkat ekonomi.

Kata Kunci: pendidikan, sains data, neural network

Abstract

Education is an important thing for every individual and determines the development of the country. Based on data from the 2015 ASEAN Yearbook, Singapore is the country with the highest education index score with a score of 0.768. Singapore not only has the best quality education system in ASEAN but also the world. With the Education Index score (EDI) of 0.692, Brunei Darussalam is in the 30th position in the world and number two in Southeast Asia. Malaysia ranks 62nd in the list of the best education in the world and third in ASEAN with an EDI score of 0.671. To reduce literacy rates, strategic points are needed to increase the literacy rate in ASEAN countries. The research method used is by using the data provided by ASEAN DSE as primary data and world bank data and UNESCO as secondary data. Data processing using the IBM SPSS Modeler. Neural Network Algorithm is used to determine the significant factors affecting the literacy rate in ASEAN. From the research, it is found that there are significant factors that affect the literacy rate in ASEAN, namely: Government budget figures in the education sector, the ratio of the number of teachers to the number of students, the ratio of the number of female students to the number of male students, the population numbers who attend primary and secondary schools based on residence, and the population who attend primary and secondary schools based on the economic level.

Keywords: literacy rates, data science, neural network

PENDAHULUAN

Pendidikan menjadi suatu hal yang penting bagi setiap orang. Bahkan pendidikan juga menjadi hal yang penting bagi negara karena pendidikan menjadi penentu bagi perkembangan negara. Semakin tinggi kualitas pendidikan di suatu negara maka dapat

dikatakan negara tersebut memiliki kualitas sumber daya yang unggul. Peningkatan kualitas SDM merupakan keharusan yang mutlak diperlukan dalam menghadapi era disrupsi.

Negara-negara ASEAN mencakup Indonesia, Malaysia, Filipina, Singapura, Thailand, Brunei, Vietnam, Laos, Myanmar dan Kamboja. Singapura dengan skor nilai Indeks Pendidikan (EDI) sebesar 0,768, tidak hanya memiliki sistem pendidikan berkualitas terbaik di ASEAN, tapi juga dunia. Brunei Darussalam dengan skor EDI sebesar 0,692 menempati posisi 30 di dunia dan nomor dua di Asia Tenggara. Malaysia menempati posisi 62 dalam daftar pendidikan terbaik di dunia dan ketiga di ASEAN dengan skor EDI 0,671. Thailand sebagai salah satu negara ASEAN yang memiliki anggaran pendidikan tertinggi, yakni 7,6% dari produk domestik bruto menempati posisi 89 di dunia dengan skor EDI sebesar 0.608. Saat ini Indonesia berada di posisi 108 di dunia dengan skor 0,603. Hanya sebanyak 44% penduduk menuntaskan pendidikan menengah, sementara 11% murid gagal menuntaskan pendidikan alias keluar dari sekolah. Filipina memiliki tingkat kegagalan murid menuntaskan sekolah yang termasuk tertinggi di dunia, yakni 24,2% dengan skor 0,610. Namun begitu sebanyak 64% penduduk menuntaskan pendidikan menengah. Vietnam yang berada di posisi 121 memiliki kualitas pendidikan yang lebih rendah ketimbang Irak dan Suriah. Saat ini Vietnam mencatat skor EDI 0,513 dan tingkat literasi penduduk dewasa sebesar 93,5%. Kamboja di peringkat 136 di dunia dengan skor 0,495. Wajah pendidikan negeri kamboja itu termasuk yang paling muram, dengan tingkat kegagalan murid sebesar 35,8% dan hanya 15,5% penduduk yang mengenyam pendidikan tingkat menengah. Laos tergolong yang paling rendah, yakni 72,7%. Setidaknya 40% penduduk belum pernah mengecap pendidikan formal. Myanmar sedang membangun kembali pendidikannya yang tertinggal. Saat ini Myanmar berada di urutan 150 di dunia dengan skor EDI 0.371. Tercatat hanya 19% penduduk Myanmar yang pernah mengecap pendidikan tingkat menengah. (Ahmad Sahroji, 2017)

Salah satu hal yang perlu dilakukan dalam rangka perbaikan kualitas pendidikan dan sumber daya manusia adalah pemberantasan buta aksara atau buta huruf dan buta berhitung. Karena buta aksara dan buta berhitung merupakan salah satu faktor yang menghambat kualitas sumber daya manusia. Dalam upaya mengatasi permasalahan strategi pemberantasan buta aksara dan buta berhitung di negara-negara ASEAN digunakan pendekatan sains data untuk memunculkan wawasan-wawasan baru yang mungkin saja belum pernah ditemukan sebelumnya terkait dengan bagaimana memberantas buta aksara dan buta berhitung di negara-negara ASEAN.

Neural Network atau Jaringan syaraf tiruan (JST) adalah paradigma pemrosesan suatu informasi yang terinspirasi oleh sistem sel saraf biologi. Neural network memiliki beberapa properti yang membuat mereka populer untuk clustering. Pertama, neural network adalah arsitektur pengolahan inheren paralel dan terdistribusi. Kedua, neural network belajar dengan menyesuaikan bobot interkoneksi dengan data, Hal ini memungkinkan neural network untuk "menormalkan" pola dan bertindak sebagai fitur (atribut) extractors untuk kelompok yang berbeda. Ketiga, neural network memproses vektor numerik dan membutuhkan pola objek untuk diwakili oleh fitur kuantitatif saja (Gorunescu, 2011). Solikhun, Mochamad Wahyudi dan Lise Pujiastuti (2019) telah memanfaatkan Neural Network dalam prediksi penduduk buta huruf di Indonesia. (Solikhun, Wahyudi, & Pujiastuti, 2019)

METODE

Pengumpulan data primer didapatkan dari bank data ASEAN DSE yang merupakan data angka literasi serta data-data pendukung lainnya. Data yang disediakan berupa data persentase pada tiap-tiap negara yang berhubungan dengan angka literasi itu sendiri. Data-data tersebut adalah:

1. *Adult literacy rate*

Data persentase populasi berusia 25-65 tahun dan lebih yang dapat membaca dan menulis. Data dikategorikan menjadi 2, yaitu data laki-laki dan perempuan.

2. *Government expenditure on edu*

Data persentase anggaran pemerintah per masing-masing negara di ASEAN untuk bidang edukasi.

3. *Last grade completion female*

Data persentase tingkat keberlangsungan siswa perempuan dari kelas awal yang diharapkan mencapai hingga kelas terakhir dari sekolah dasar pada tahun ajaran tertentu, termasuk siswa yang mengulang kelas.

4. *Last grade completion male*

Data persentase tingkat keberlangsungan siswa laki-laki dari kelas awal yang diharapkan mencapai hingga kelas terakhir dari sekolah dasar pada tahun ajaran tertentu, termasuk siswa yang mengulang kelas.

5. *Nett primary enroll rate female*

Data persentase tingkat pendaftaran di sekolah dasar siswa perempuan berdasarkan populasi dari kelompok usia di sekolah resmi.

6. *Nett primary enroll rate male*

Data persentase tingkat pendaftaran di sekolah dasar siswa laki-laki berdasarkan populasi dari kelompok usia di sekolah resmi.

7. *Primary attendance*

Data persentase populasi yang bersekolah di sekolah dasar berdasarkan tempat tinggal dan kelas ekonomi.

8. *Second attendance*

Data persentase populasi yang bersekolah di sekolah menengah berdasarkan tempat tinggal dan kelas ekonomi.

9. *Primary completion rate female*

Data persentase tingkat kelulusan sekolah dasar siswa perempuan dihitung berdasarkan jumlah siswa yang ada di tingkat akhir dikurangi jumlah siswa yang mengulang kelas, kemudian dibagi dengan jumlah keseluruhan siswa dari tingkat awal hingga tingkat akhir.

10. *Primary completion rate male*

Data persentase tingkat kelulusan sekolah dasar siswa laki-laki dihitung berdasarkan jumlah siswa yang ada di tingkat akhir dikurangi jumlah siswa yang mengulang kelas, kemudian dibagi dengan jumlah keseluruhan siswa dari tingkat awal hingga tingkat akhir.

11. *Pupil-teacher ratio*

Data persentase rasio jumlah murid dibandingkan dengan guru di sekolah. Data dikategorikan menjadi 2, yaitu data sekolah dasar dan sekolah menengah.

12. *Ratio girls to boys*

Data persentase rasio jumlah anak perempuan dibandingkan dengan anak laki-laki di sekolah. Data dikategorikan menjadi 2, yaitu data sekolah dasar dan sekolah menengah.

13. *Youth literacy rate female*

Data persentase populasi perempuan berusia 15-24 tahun yang dapat membaca & menulis.

14. *Youth literacy rate male*

Merupakan data persentase populasi laki-laki berusia 15-24 tahun yang dapat membaca & menulis.

Data-data tersebut tersebut dapat diakses secara legal pada link sebagai berikut: <https://aseandse.org/datasets>.

Data sekunder dibutuhkan untuk menjawab permasalahan edukasi ditinjau dari sisi ekonomi tentang bagaimana mengurangi angka buta huruf di negara-negara ASEAN.

Variabel yang menunjukkan angka literasi serta populasi berdasarkan tingkat ekonomi dan tempat tinggal diperoleh dari bank data UNESCO yang secara legal menyediakan data literasi serta data tingkat ekonomi dan tempat tinggal. Data yang disediakan berupa data persentase pada tiap-tiap negara. Data-data tersebut adalah:

1. *Adult literacy rate male*

Data persentase populasi laki-laki berusia 25-65 tahun dan lebih yang dapat membaca & menulis.

2. *Adult literacy rate female*

Data persentase populasi perempuan berusia 25-65 tahun dan lebih yang dapat membaca & menulis.

3. *Youth literacy rate male*

Data persentase populasi laki-laki berusia 15-24 tahun yang dapat membaca & menulis.

4. *Youth literacy rate female*

Data persentase populasi perempuan berusia 15-24 tahun yang dapat membaca & menulis.

5. *Primary attendance*

Data persentase populasi yang bersekolah di sekolah dasar berdasarkan tempat tinggal dan kelas ekonomi.

6. *Second attendance*

Data persentase populasi yang bersekolah di sekolah menengah berdasarkan tempat tinggal dan kelas ekonomi.

Data-data tersebut tersebut dapat diakses secara legal pada link sebagai berikut:

http://data.uis.unesco.org/Index.aspx?DataSetCode=EDULIT_DS&popupcustomise=true&lang=en

Data primer dan data sekunder dengan rentang waktu tahun 2004 sampai dengan tahun 2019 dengan cakupan data negara-negara di wilayah ASEAN. Dengan demikian, data ini layak untuk dijadikan model untuk menemukan signifikan factor yang menjadi penentu untuk mengurangi angka buta huruf di negara-negara ASEAN.

HASIL DAN PEMBAHASAN

Data input berasal dari data yang didapatkan dari data primer dan sekunder website World Bank dan data UNESCO kemudian di ekstrak ke lokal database dan ditarik kembali menggunakan koneksi ODBC dan melalui tools IBM SPSS untuk diubah menjadi Tabel *Data Preparation* dengan menampilkan faktor-faktor yang dibutuhkan dalam

permasalahan. File tersebut berisikan kumpulan data yang akan diolah untuk memprediksi angka literasi pada negara-Negara di ASEAN. Pada kasus prediksi angka literasi pada negara-negara di ASEAN, terdapat berbagai jenis data yang diolah menggunakan logika algoritma neural network untuk menghasilkan sebuah prediksi dari angka literasi pada tahun 2020-2023. Data-data yang digunakan pada penelitian ini terdiri dari:

1. Data anggaran edukasi pemerintah pada masing-masing negara di ASEAN tahun 2004-2009.
2. Data literasi populasi yang berusia 25-65 tahun pada masing-masing negara di ASEAN tahun 2004-2009.
3. Data literasi populasi yang berusia 15-24 tahun pada masing-masing negara di ASEAN tahun 2004-2009.
4. Data tingkat pendaftaran sekolah dasar pada masing-masing negara di ASEAN tahun 2004-2009.
5. Data tingkat kelulusan sekolah dasar pada masing-masing negara di ASEAN tahun 2004-2009.
6. Data tingkat keberlangsungan siswa/i dari kelas awal di sekolah dasar pada masing-masing negara di ASEAN tahun 2004-2009.
7. Data rasio jumlah anak perempuan dibandingkan dengan jumlah anak laki-laki pada masing-masing negara di ASEAN tahun 2004-2009.
8. Data rasio jumlah murid dibandingkan dengan jumlah guru pada masing-masing negara di ASEAN tahun 2004-2009.
9. Data populasi yang bersekolah di sekolah dasar dan menengah berdasarkan tempat tinggal pada masing-masing negara di ASEAN tahun 2004-2009.
10. Data populasi yang bersekolah di sekolah dasar dan menengah berdasarkan tingkat ekonomi pada masing-masing negara di ASEAN tahun 2004-2009.

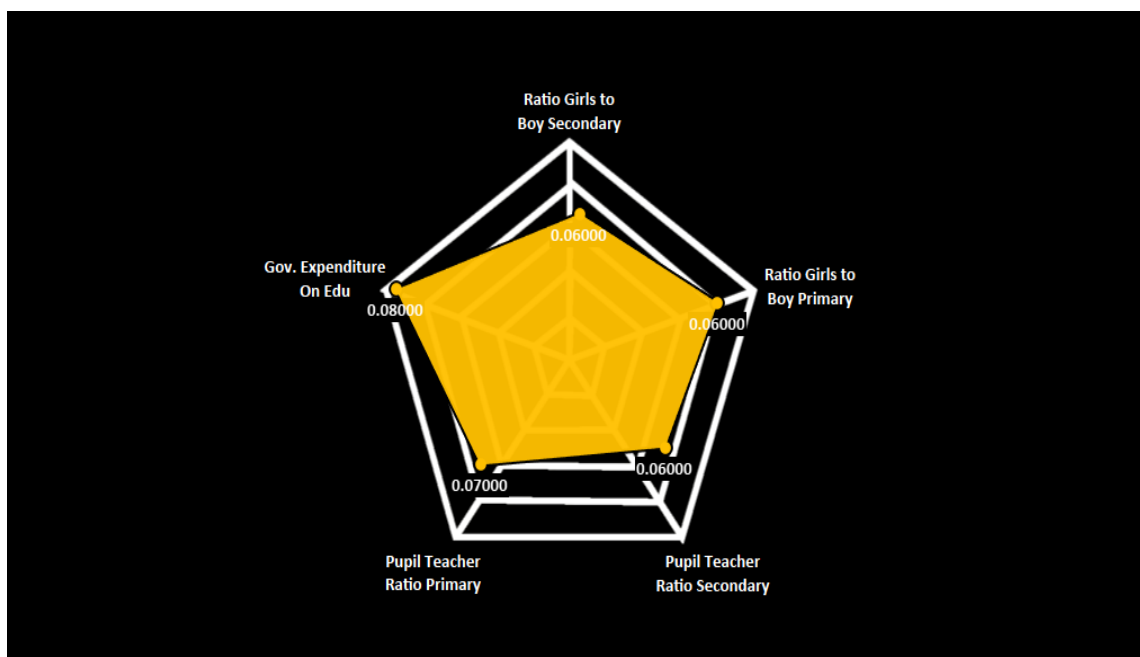
Output dari sistem informasi strategi berupa dashboard informasi yang memuat informasi tentang tingkat literasi berdasarkan tiap-tiap negara, prediksi angka literasi dari tahun ke tahun dan faktor signifikan yang mempengaruhi angka literasi. Output memuat informasi signifikan faktor apa saja yang berperan untuk mengurangi angka buta huruf. Kemudian Output juga memuat informasi persen penurunan angka buta huruf dari tahun ke tahun. Fungsi tambahan dari output adalah bisa diakses oleh publik sebagai referensi dalam hal kasus buta huruf.

Produk yang dimaksud yaitu berupa dashboard informasi yang akan di publikasi kepada pihak ASEAN DSE dan untuk insight analisis penelitian selanjutnya mengenai topik terkait.

Produk diolah menggunakan:

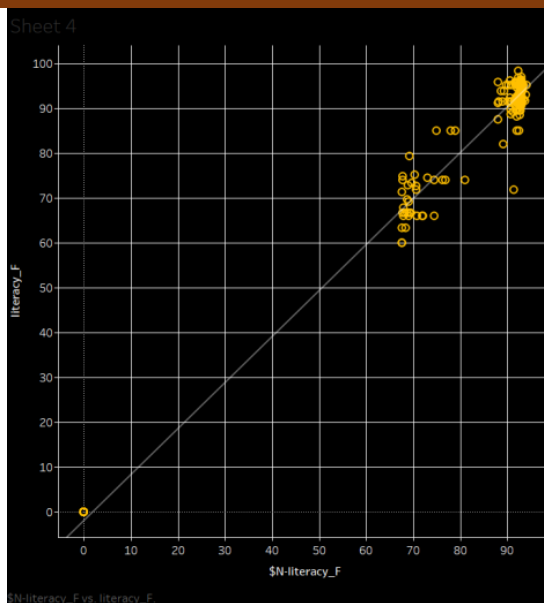
1. MySQL sebagai database server
2. IBM SPSS Modeler for student sebagai aplikasi penambangan data, teks analitik, membuat model, dan menemukan pola yang memiliki antar muka visual yang memungkinkan pengguna untuk memanfaatkan algoritma statistik dan data mining tanpa pemrograman.
3. *Tableau Desktop for student* sebagai *tools* yang digunakan untuk menganalisa / menggambarkan suatu kumpulan data untuk disajikan dalam bentuk grafik yang menarik.

Produk yang telah selesai dirancang merupakan hasil dari kumpulan data yang telah diolah menggunakan beberapa tools yang telah disebutkan sebelumnya. Hasil dari produk tersebut untuk kedepannya akan dapat diakses oleh semua orang setelah dipublikasi oleh pihak ASEAN DSE. Berikut adalah beberapa rancangan tampilan yang menunjukkan dashboard informasi. Tampilan ini memberikan informasi yaitu faktor-faktor yang secara signifikan mempengaruhi penurunan angka buta huruf serta korelasi data target dengan data hasil algoritma.

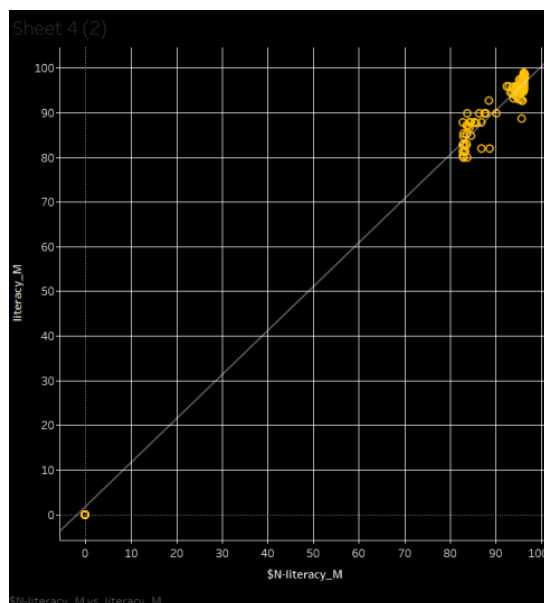


Gambar 1. Significant Factor Literacy Rate

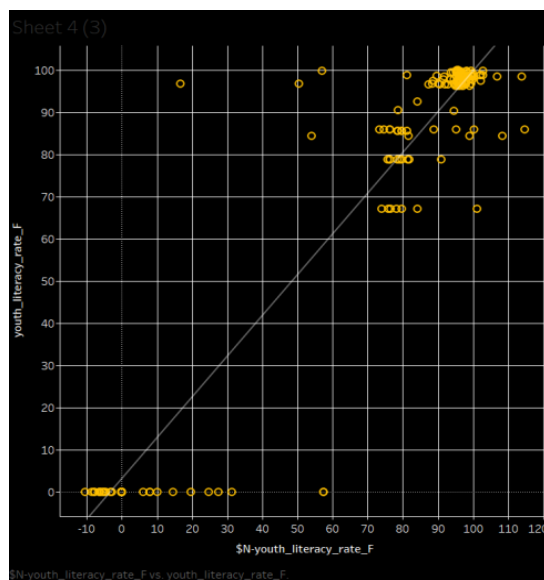
Setelah data diproses di IBM SPSS Modeler for Student, dapat diketahui bahwa faktor yang mempengaruhi angka literasi di ASEAN antara lain: Angka anggaran pemerintah pada bidang pendidikan, rasio jumlah guru dengan jumlah murid laki-laki di sekolah dasar, rasio jumlah guru dengan jumlah murid laki-laki di sekolah menengah, rasio jumlah murid perempuan dibandingkan dengan jumlah murid laki-laki di sekolah menengah, dan rasio jumlah murid perempuan dibandingkan dengan jumlah murid laki-laki di sekolah dasar.



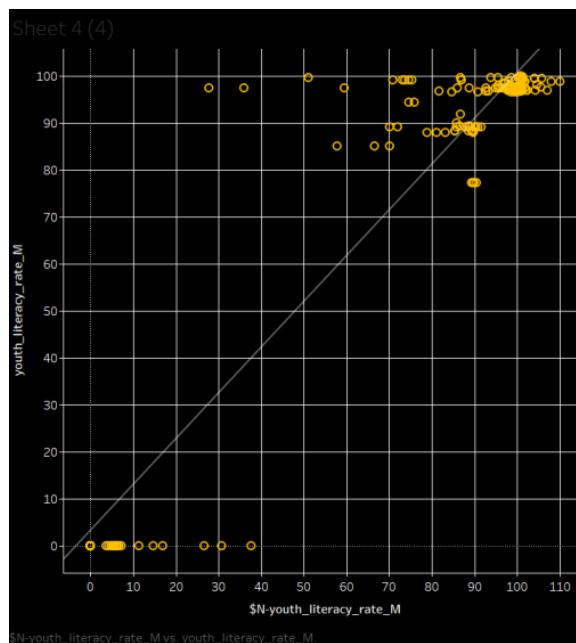
Gambar 2. Analisa Korelasi Target 1



Gambar 3. Analisa Korelasi Target 2



Gambar 4. Analisa Korelasi Target 3

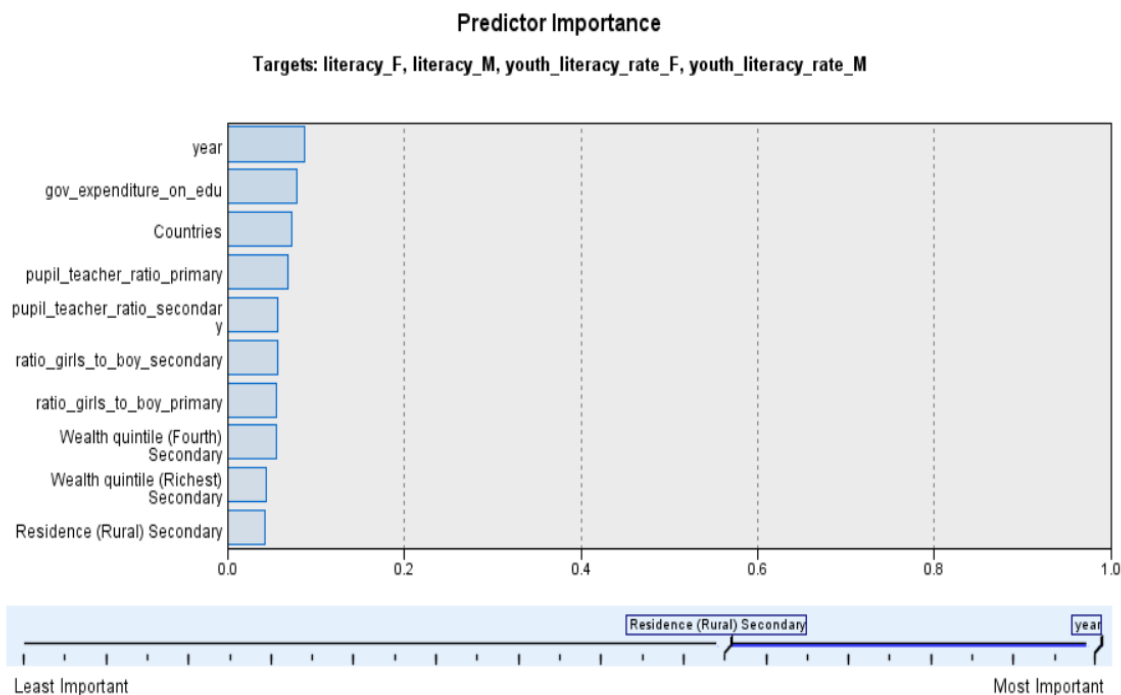


Gambar 5. Analisa Korelasi Target 4

$R squared$ merupakan angka yang berkisar antara 0 sampai 1 yang mengindikasikan besarnya kombinasi variabel independen secara-sama mempengaruhi nilai variabel dependen. Semakin mendekati angka satu, model yang dikeluarkan oleh regresi tersebut akan semakin baik. Jika dilihat dari keempat gambar analisa korelasi diatas, dapat disimpulkan bahwa keempatnya memiliki $R Squared$ yang mendekati angka 1, maka dapat dikatakan bahwa model yang dihasilkan sudah dapat dikatakan baik.

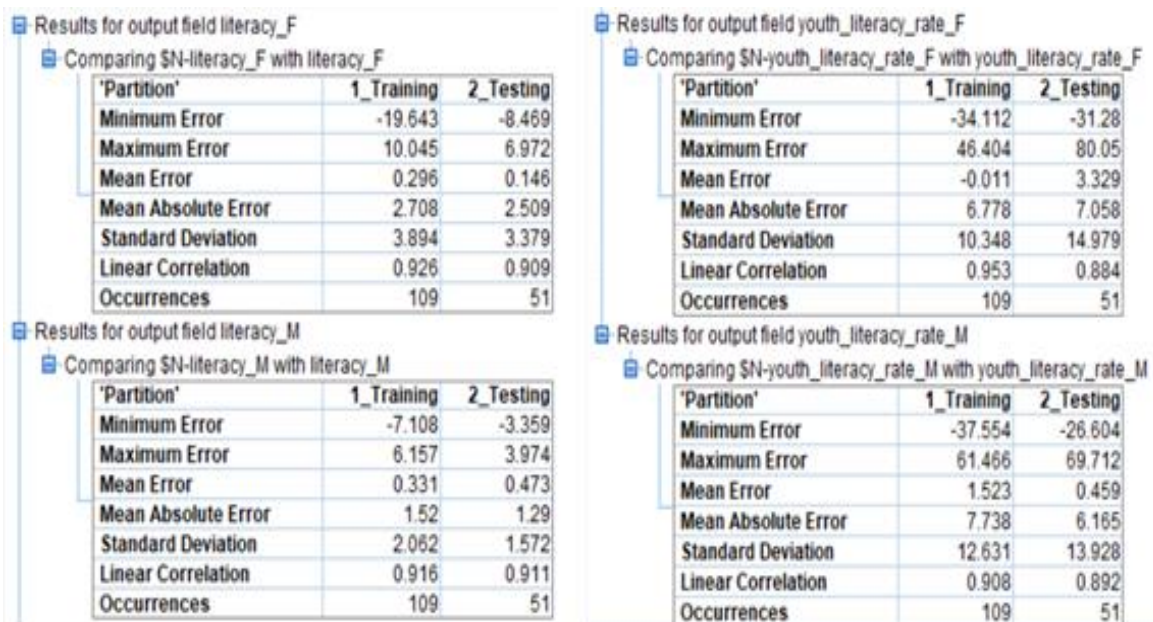
Data scientist melakukan pemodelan data dengan menguji model mana yang paling baik untuk permasalahan ini. Merujuk dari business understanding, maka permasalahan ini termasuk problem estimasi dan prediksi. Langkah selanjutnya pada fase ini adalah menetapkan algoritma klasifikasi dan prediksi. Algoritma klasifikasi dipilih algoritma Neural network, dimana algoritma ini digunakan untuk menyelesaikan permasalahan klasifikasi dengan target numerik. Kemudian dari proses modeling, diperoleh signifikan faktor sebagai gambar 6.

Hasil model menyatakan signifikan faktor yang mempengaruhi angka literasi adalah anggaran pemerintah pada bidang pendidikan menjadi *predictor* paling penting dengan bobot 0.08 dan peringkat kedua signifikan faktor adalah ratio jumlah guru dengan jumlah murid dengan bobot 0.07.



Gambar 6. Signifikan Faktor Hasil Proses Modeling

Fase evaluation adalah menilai seberapa *accuracy model*. Neural network adalah model prediksi yang menghasilkan prediksi literasi. Oleh karena hasilnya adalah angka jumlah, maka untuk mengevaluasi model, bisa dilihat dari nilai *linear correlation*. Model dinilai semakin baik jika nilainya *linear correlation* mendekati angka 1. Gambar dibawah ini adalah gambar output node analisis yang menampilkan *linear correlation* = 1 baik pada data training dan data testing.



Gambar 7. Analisis *Linear Correlation*

SIMPULAN

Dari hasil perancangan yang telah diuraikan pada Strategi Pemberantasan Buta Aksara Dan Buta Berhitung di Negara-negara Asean Menggunakan Pendekatan Data Science dapat disimpulkan sebagai berikut:

1. Output dari hasil perancangan ini berupa dashboard information yang memuat informasi tentang faktor-faktor signifikan yang mempengaruhi angka melek huruf / literasi di ASEAN dan analisa korelasi dari target yang telah ditentukan dengan data hasil algoritma.
2. Faktor yang mempengaruhi angka literasi antara lain: Angka anggaran pemerintah pada bidang pendidikan, rasio jumlah guru dengan jumlah murid laki-laki di sekolah dasar, rasio jumlah guru dengan jumlah murid laki-laki di sekolah menengah, rasio jumlah murid perempuan dibandingkan dengan jumlah murid laki-laki di sekolah menengah, dan rasio jumlah murid perempuan dibandingkan dengan jumlah murid laki-laki di sekolah dasar.
3. Strategi yang dapat dilakukan dari beberapa signifikan faktor yang telah di dapat antara lain: Pemerintah lebih memperhatikan atau mempertahankan anggaran dalam bidang pendidikan untuk memberikan ruang ataupun fasilitas bagi warganya untuk mengenyam pendidikan sehingga angka buta huruf dan buta berhitung semakin kecil setiap tahunnya. Memperbesar angka rasio jumlah guru untuk menekan angka buta berhitung dan buta huruf.

REFERENSI

- Ahmad Sahroji. (2017). Daftar Negara ASEAN dengan Peringkat Pendidikan Tertinggi. Retrieved from okezone.com website:
<https://news.okezone.com/read/2017/11/24/18/1820178/daftar-negara-asean-dengan-peringkat-pendidikan-tertinggi>
- Gorunescu, F. (2011). *Data Mining Concepts, Models and Techniques*. USA: Springer.
- Solikhun, Wahyudi, M., & Pujiastuti, L. (2019). Jaringan Syaraf Tiruan Memprediksi Penduduk Buta Huruf Menurut Provinsi. In indoms (Ed.), *Seminar Nasional Matematika dan Terapan* (Vol. 1, pp. 271–278). Pematang Siantar: STMIK Tunas Bangsa.