**Klasifikasi Lama Waktu Gejala Vaksin Covid-19 dengan Metode *K-Nearest Neighbor*, *Random Forest* dan *Support Vector Machine***

Studi Kasus Pegawai Rumah Sakit Di Rumah Sakit Rujukan Nasional Indonesia

*(Classification of the Duration of Covid-19 Vaccine Symptoms with the K-Nearest Neighbor, Random Forest and Support Vector Machine Methods)*

Riofebri Prasetia1

*1Politeknik Statistika Sekolah Tinggi Ilmu Statistik*

E-mail: [221911192@stis.ac.id](mailto:abcd@ggg.com)

**ABSTRAK**

Pada akhir tahun 2019 virus penyakit covi-19 muncul di Wuhan, China. Virus ini memiliki kemampuan menyebar yang sangat cepat sehingga tidak lama kemudian dilaporkan sebanyak lebih dari 100 negara terkena wabah ini termasuk Indonesia. Berbagai macam kebijakan pemerintah telah dilakukan untuk memutus rantai penyebaran virus covid-19, satu diantaranya yaitu dengan vaksinasi gratis. Banyak opini dengan sentimen positif oleh masyarakat tentang vaksinasi sehingga menjadi kemudahan dari pemerintah untuk mengajak masyarakat ikut vaksinasi. Sementara itu, dari tenaga medis perlu tenaga ekstra dan menguras sangat banyak waktu ketika berhadapan dengan penderita covid-19 dan masyarakat yang di vaksin. Karena vaksin punya gejala bagi masyarakat sudah vaksinasi yang tidak dapat dipastikan sehingga perlu dibuat model klasifikasi untuk melakukan keputusan dengan cepat dan efisien ketika berhadapan dengan banyak masyarakat yang di vaksin. Dalam hal ini dilakukan data mining klasifikasi berdasarkan lama gejala dengan metode *K-Nearest Neighbor*, *Random Forest* dan *Support Vector Machine*. Hasilnya perbedaan hasil ke-tiga model tidak signifikan yaitu rata-rata *accuracy* sebesar 58%. Ketika dilakukan evaluasi model, didapat model *Random Forest* mendapat hasil yang lebih baik dibanding ke-dua model lainnya.

**Kata kunci**: Vaksinasi, waktu gejala, covid-19

***ABSTRACT***

*At the end of 2019, the COVID-19 virus emerged in Wuhan, China. This virus has the ability to spread very quickly, so it was not long before it was reported that more than 100 countries were affected by the outbreak, including Indonesia. Various policies have been carried out by the government to break the chain of the spread of the Covid-19 virus, one of which is free vaccination. There are many opinions with positive sentiments from the public about vaccination, making it easier for the government to invite people to participate in vaccination. Meanwhile, medical personnel need extra energy and consume a lot of time when dealing with COVID-19 sufferers and people who are vaccinated. Because vaccines have symptoms for people who have been vaccinated which cannot be confirmed, it is necessary to create a classification model to make decisions quickly and efficiently when dealing with many people who are vaccinated. In this case, classification data mining is carried out based on the duration of symptoms using the K-Nearest Neighbor, Random Forest and Support Vector Machine methods. The result is that the difference in the results of the three models is not significant, namely the average accuracy of 58%. When evaluating the model, it was found that the Random Forest model got better results than the other two models.*

***Keywords****: vaccinations, symptom timing, covid-19*

**PENDAHULUAN**

Pada akhir tahun 2019, Dunia dikejutkan dengan kemunculan penyakit Covid-19 yang terjadi di Wuhan, Cina. Menurut *Shereen et al. (2020)*, Covid-19 merupakan suatu wabah infeksi virus dan patogen yang disebabkan oleh Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-Cov-2). Pada tanggal 12 Januari 2020, The National Health Commission of China merilis bahwa hal tersebut merupakan bagian dari virus pneumonia. Dugaan sementara yaitu orang yang terinfeksi merupakan pengunjung pasar makanan laut atau tempat hewan hidup dijual. Akan tetapi hal ini terbantah dengan hasil penyelidikan lebih lanjut yang mengungkapkan bahwa beberapa orang yang terinfeksi tercatat tidak mengunjungi tempat tersebut. Penyebaran virus ini sangat cepat hingga tidak lama kemudian dilaporkan lebih dari 100 negara di dunia terkena wabah termasuk Indonesia.

Wabah ini khususnya di Indonesia sangat berdampak tidak hanya di bidang kesehatan, tapi juga di bidang ekonomi, pendidikan dan masih banyak lagi. Menurut *Permatasari (2021)*, Pembatasan Sosial Berskala Besar (PSBB ) merupakan kebijakan alternatif pemerintah untuk mencegah persebaran Covid-19. Selanjutnya disusul dengan Pemberlakuan Pembatasan Kegiatan Masyarakat (PPKM) yang ditetapkan untuk beberapa daerah di Indonesia. Adapun satu di antara kebijakan yang masih berjalan hingga saat ini yaitu pemberian vaksinasi ke masyarakat dengan alasan pembangunan kekebalan komunitas (Farisa, 2020).

Beberapa opini yang telah diberikan masyarakat tentang vaksinasi. Berdasarkan media sosial Twitter, Masyarakat lebih banyak memberikan opini yang bersentimen positif dibanding negatif (Rachman et al., 2020). Kemudian berdasarkan survei nasional penerimaan vaksin covid-19 dalam Rokom (2021), Sebagian besar masyarakat Indonesia bersedia menerima vaksin covid-19. Akan tetapi sekitar 7% dari masyarakat yang menolak, memiliki alasan keamanan, kehalalan dan efektivitas vaksin. Dalam hal ini dapat dimungkinkan untuk mengajak masyarakat untuk vaksinasi Covid-19.

Menurut Bramasta (2021), selain Badan Pengawas Obat dan Makanan (BPOM) menyatakan keamanan vaksin covid-19, tetap ada beberapa upaya pemerintah agar masyarakat tidak cemas untuk vaksinasi. Pemerintah memberikan edukasi ke masyarakat tentang vaksin dan meyakinkan masyarakat bahwa tidak ada kerugian bila mengikuti vaksinasi. Walaupun tetap ada kemungkinan orang yang sudah vaksin bisa terinfeksi virus covid-19, vaksin ini membentuk antibodi dan mengurangi gejala infeksi virus covid-19 dibanding dengan gejala pada orang yang belum di vaksinasi.

Hingga saat ini pemberian vaksin mencapai lebih dari 200 juta suntikan dimana sebanyak 123% suntikan dosis pertama, 77% suntikan dosis kedua dan 1% untuk suntikan dosis ke tiga (Rokom, 2021). Dalam hal ini diharapkan sebagian besar populasi di Indonesia mendapat dosis pertama akhir desember tahun 2021 (COVID-19, 2021). Adapun prioritas yang memperoleh vaksin berdasarkan Bab III pasal 8 PMK Nomor 84/2020 dimana kelompok prioritas pertama yaitu tenaga kesehatan, asisten tenaga kesehatan, tenaga penunjang yang bekerja pada fasilitas kesehatan (“Urutan Daftar Prioritas Penerima Vaksin Covid-19,” 2020).

Walaupun vaksin telah dilakukan bukan berarti masalah selesai. Dalam (Rokom, 2021), butuh waktu 28 hari untuk membentuk kekebalan setelah penyuntikan kedua. Adapun vaksinasi ini tidak dipastikan melindungi diri dari penyakit covid-19, tetapi bisa mengurangi paparan atau infeksi yang disebabkan virus covid-19. Dalam hal ini, masyarakat yang sudah di vaksinasi tetap harus menjaga protokol kesehatan yang berlaku karena ada kemungkinan bisa terpapar virus covid-19. Selain itu menurut hasil jawaban pertanyaan terkait covid-19 pada Kementerian Kesehatan (2021), terdapat efek samping ringan setelah vaksinasi dan tidak terjadi pada semua orang. Efek samping ringan yang terjadi seperti demam, nyeri otot, dan ruam-ruam pada area suntikan. Dikatakan juga efek samping ini bersifat sementara.

Dengan banyaknya kasus covid-19 tentunya membawa banyak beban yang diberikan kepada tenaga medis untuk melayani penderita covid-19. Menurut Handayani et al. (2020), durasi pekerjaan yang lama, persediaan alat pelindung diri yang kurang, serta tekanan yang diberikan oleh pembawa virus yang sebelumnya belum pernah terjadi merupakan sebagian faktor penyebab stres pada tenaga medis. Belum lagi kekhawatiran dengan virus tidak terlihat yang mungkin bisa menginfeksi tenaga medis. Hal ini lah yang menjadi alasan kenapa tenaga kesehatan di prioritas pertama, karena seringnya keterlibatan dengan Covid-19 oleh tenaga kesehatan sehingga, diharapkan tenaga kesehatan mendapat keamanan terlebih dahulu (Ramadhan, 2021)

Dengan kondisi demikian, tentunya dalam penanganan kasus covid-19 perlu adanya keputusan yang cepat, tepat dan efisien tidak hanya kepada penderita covid-19 tapi juga terhadap tenaga medis atau staff di rumah sakit yang bertugas. Terutama ketika menangani penderita efek samping dari vaksinasi. Untuk itu perlu adanya klasifikasi lama gejala berdasarkan model gejala yang timbul setelah dilakukan vaksinasi agar bisa dilakukan penanganan yang tepat kepada penderita covid-19. Tentunya dalam metode klasifikasi yang digunakan perlu dilakukan pengujian untuk memilih metode terbaik menentukan model “lama waktu gejala” setelah vaksin.

Berdasarkan yang sudah di paparkan, peneliti tertarik untuk mengetahui model klasifikasi “lama gejala setelah vaksinasi” dengan pendekatan data mining. Dan juga dilakukan perbandingan metode klasifikasi antara metode *K-Nearest Neighbor*, *Random Forest* dan *Support Vector Machine*.

**METODE**

1. Sumber data dan deskripsi atribut

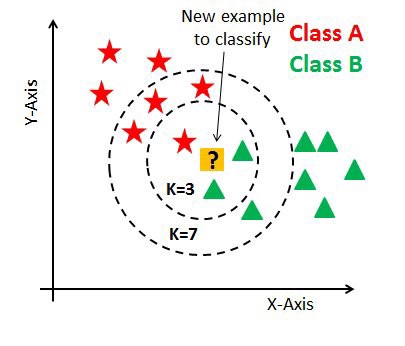
Data yang digunakan berasal dari survei efek samping vaksin Covid-19 pada staff rumah sakit di rumah sakit rujukan nasional Indonesia. Data berjumlah 840 responden yang mengisi dan terdapat sebanyak 22 pertanyaan.

Table 1. Tabel Variabel dan Posisi Variabel

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **No** | | **Nama** | | **Nilai** | |
| 1 | Jenis Kelamin | | * Laki-laki * Perempuan | |
| 2 | Usia | | * < 20 tahun * 20-25 tahun * 26-30 tahun * 31-35 tahun * 36-40 tahun * 41-45 tahun * 46-50 tahun * 51-55 tahun * 56-60 tahun * > 60 tahun | |
| 3 | Profesi | | * Bidan * Perawat * Dokter medis * Dokter Spesialis * Staff non medis | |
| 4 | Pendidikan | | * SMP * SMA * Sarjana/S1 * Master/S2 | |
| 5 | Tempat tinggal | | * Pusat Kota * Pinggir Kota | |
| 6 | Pembengkakan | | * Ya * Tidak | |
| 7 | Kemerahan | | * Ya * Tidak | |
| 8 | Gatal | | * Ya * Tidak | |
| 9 | Demam | | * Ya * Tidak | |
| 10 | Sakit Kepala | | * Ya * Tidak | |
| 11 | Nyeri Otot | | * Ya * Tidak | |
| 12 | Kelelahan | | * Ya * Tidak | |
| 13 | Batuk | | * Ya * Tidak | |
| 14 | Diare | | * Ya * Tidak | |
| 15 | Mual dan Muntah | | * Ya * Tidak | |
| 16 | Sesak Napas | | * Ya * Tidak | |
| 17 | Nyeri Sendi | | * Ya * Tidak | |
| 18 | Pingsan | | * Ya * Tidak | |
| 19 | Reaksi Anafilaksis | | * Ya * Tidak | |
| 20 | Perasaan geli | | * Ya * Tidak | |
| 21 | Pembengkakan kelenjar getah bening | | * Ya * Tidak | |
| 22 | Waktu Gejala | | * Tidak ada * < 24 jam * 24-72 jam * > 72 jam | |

1. *K-Nearest Neighbor*

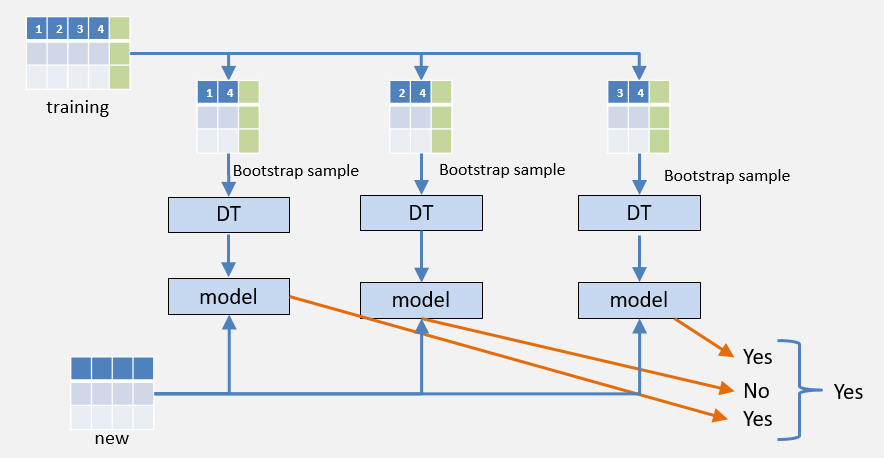
*K-Nearest Neighbor* merupakan metode pengklasifikasian dengan tetangga terdekat berdasarkan pembelajaran dengan analogi. “Nearest” atau Kedekatan didefinisikan dengan jarak metrik, seperti jarak *Euclidian* (Han & Kamber, 2006)*. Instance* di klasifikasi berdasarkan jarak dengan k tetangga terdekat dimana k merupakan parameter bebas.



Gambar 1. Ilustrasi Metode K-Nearest Neighbor Sumber: medium.com

1. *Random Forest*

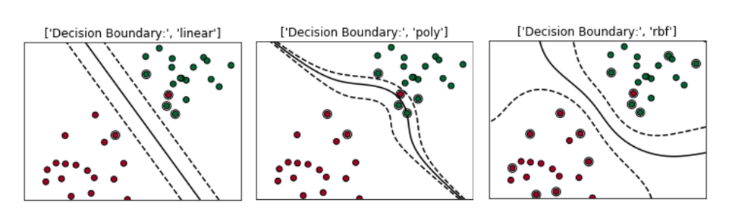
*Random Forest* merupakan satu diantara pembelajaran ensemble. Berdasarkan Speiser et al. (2019), metode ini menggunakan model sederhana yang menggunakan pemisahan biner pada variabel prediksi untuk menentukan prediksi hasil. Sistem kerja *Random Forest* sendiri yaitu mengambil subset *feature* secara acak sehingga secara sistematis menghindari korelasi dan meningkatkan kinerja model (Singh, 2019).



Gambar 2. Ilustrasi Metode Random Forest Sumber: medium.com

1. *Support Vector Machine*

Berdasarkan (Han & Kamber, 2006), *Support Vector Machine* merupakan metode untuk klasifikasi data linier dan non linier menggunakan pemetaan non linier untuk mengubah *training dataset* menjadi dimensi yang lebih tinggi. Pada dimensi baru dicari hyperplane pemisah optimal linier. Hyperplane ini menggunakan vektor dukungan dan margin yang didefinisikan vektor dukungan.



Gambar 3. Ilustrasi Metode SVM dengan Jenis Decision Boundary yang digunakan. Sumber: towardsdatascience.com

1. Alur Data Mining

Data Mining

Preprocessing

Seleksi

Transformasi

Interpretasi dan Evaluasi

Gambar 4. Diagram alur Data Mining.

* Seleksi

Mengubah dataset menjadi format yang sesuai untuk analisis. Pada tahap ini akan di seleksi atribut. Akan di lakukan reduksi dimensi dan sub setting data jika dibutuhkan.

* Preprocessing

Merupakan tahap pembersihan data dimana informasi yang tidak dibutuhkan dibuang. Data di setting ulang untuk memastikan format konsisten.

* Data Mining

Pada tahap ini dilakukan pengujian model yang mana label yang diklasifikasikan ialah “Waktu gejala”. Akan dilakukan pengujian model dimana proporsi yang ditentukan ialah 67% untung *training set* dan 33% untuk *test set.* Setelah dilakukan pengujian model, selanjutnya dilakukan evaluasi dan validasi hasil. Untuk evaluasi hasil menggunakan confusion matrix dan hasil klasifikasi berupa *precision, recall/sensitivity*, *F1-score* dimana dengan ini bisa dihitung *accuracy.* Kemudian pada validasi hasil digunakan *cross validation.*

1. *Confusion Matrix*

*Confusion matrix* merupakan gambaran pengelompokan kelas yang terdiri dari kelas sebenarnya/*retrieved* dengan kelas yang diprediksi/*relevant*. Dengan ini bisa digambarkan jumlah data/*instance* yang termasuk kelompok kelas berikut:

Table 2. Configuration Matrix

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Predicted Class  */Relevant* | True/Actual Class/*Retrieved* | | | |
| P/A | Positive | Negative |  |
| True | True Positif  (TP) | False Positive  (FP) | P’ |
| False | False Negatif  (FN) | True Negative  (TN) | N’ |
|  | P | N | All |

1. *Precision*

Presisi adalah persentase yang relevan/*relevant* dengan tanggapan “benar” atau *True Positive*. Secara formal didefinisikan sebagai:

1. *Recall/sensitivity*

*Recall* merupakan persentase yang relevan/*relevant* dengan tanggapan “benar” yang pada kenyataannya diambil/*retrieved.*  Secara formal dapat didefinisikan sebagai:

1. *F1-score*

F1-score merupakan perbandingan rata-rata harmonik dari *recall* dan *precision.* Secara umum dapat didefinisikan sebagai berikut:

1. *Accuracy*

Ketepatan prediksi baik yang bernilai *positive* maupun bernilai *negative*, diukur dari persentase *accuracy.* Secara umum dapat didefinisikan sebagai berikut:

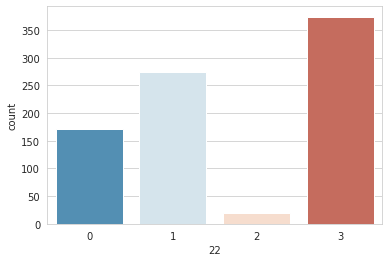
1. *Cross validation*

*Cross validation* merupakan satu di antara metode resampling data sampel untuk menilai generalisasi kemampuan model *predictive* dan mencegah overfitting. Menurut Berrar (2018), *cross validation* Biasanya digunakan untuk memberikan perkiraan kinerja model dengan menyetel parameter model yang diterapkan beberapa kali pada nilai parameter yang berbeda, dan parameter yang dicoba bisa meminimalkan kesalahan yang kemudian digunakan untuk membangun model akhir. Dengan demikian *cross validation* mengatasi masalah overfitting.

* Interpretasi dan evaluasi

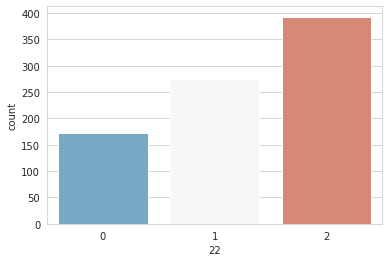
Setelah didapat hasil terbaik, maka hasil model akan di interpretasi dan evaluasi. Adapun interpretasi dan evaluasi ini difokuskan pada *accuracy, precision, sensitivity* serta akurasi model. Serta bagaimana model yang terbaik dari berbagai percobaan yang dilakukan sesuai dengan kondisi data.

**HASIL DAN PEMBAHASAN**



Gambar 5. Diagram batang frekuensi kolom pertanyaan waktu gejala

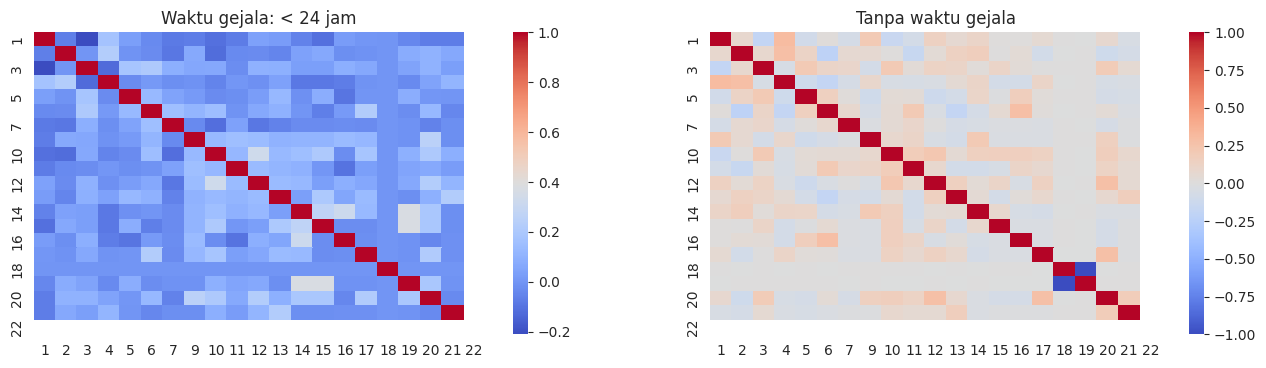
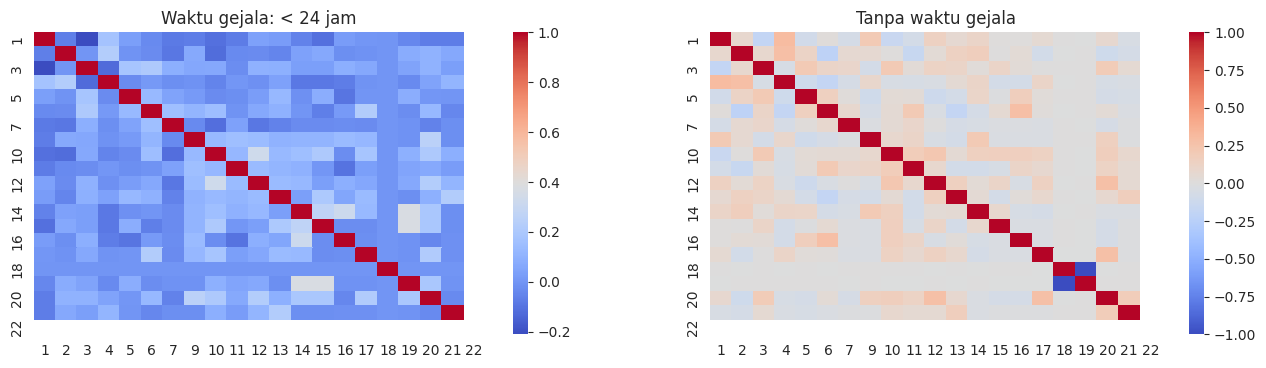
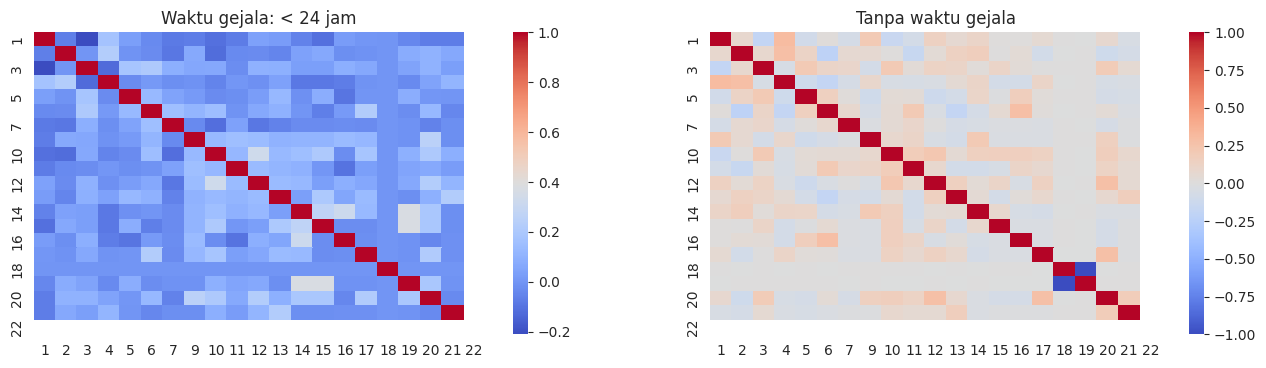
Pada kolom pertanyaan 22 yaitu waktu gejala terjadi extreme imbalance pada opsi jawaban 2 yakni 24 – 72 jam. Sehingga dalam hal ini dilakukan agregasi baris sehingga dijadikan 3 kelompok nilai klasifikasi yakni waktu gejala: “tidak ada”, “< 24 jam” dan “> 24 jam”. Sehingga menjadi:



Gambar 6. Diagram batang frekuensi kolom pertanyaan waktu gejala setelah agregasi baris

Jika diperhatikan terjadi imbalance dimana jumlah responden yang menjawab lama gejala: “tidak ada” lebih kecil kurang dari 50% jumlah responden yang menjawab “> 24 jam”. Selanjutnya di lakukan drop pada atribut pertanyaan “gatal” karena jawaban semua respon adalah “tidak”, sehingga atribut ini tidak bisa menjadi *feature* untuk membantu klasifikasi.

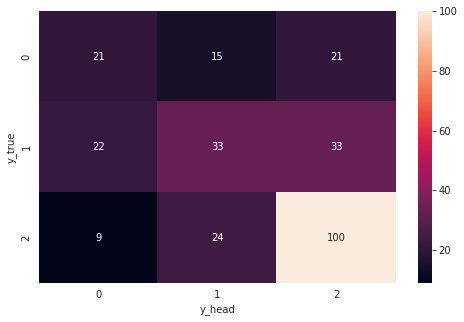
Kemudian dilihat bagaimana pengaruh *feature* dalam mengklasifikasikan label “Waktu gejala” serta perbandingannya pada nilai klasifikasi yang lain seperti pada gambar



Gambar 7. Korelasi perbedaan antar feature terhadap label

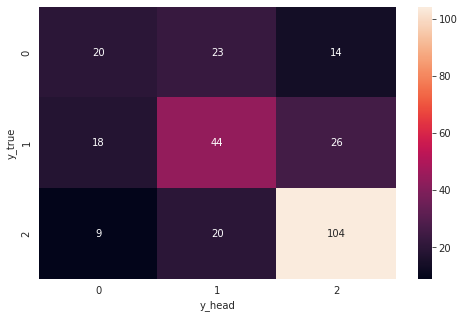
Hingga saat ini masih belum terdapat atribut antar *feature* yang memiliki korelasi yang tinggi. Sehingga belum ada perlakuan khusus terhadap *features.*

Kemudian didapat hasil dari “*cross validation”* dimana ketika dilakukan uji dengan data set yang sama, menghasilkan klasifikasi model dengan confusion matrix:



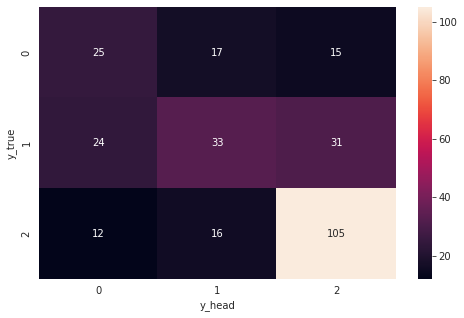
Gambar 8. Confusion Matrix pada model K-Nearest Neighbor

Dalam model *K-Nearest Neighbor* bisa kita lihat bahwa terdapat error dimana terdapat 21 orang yang diprediksi tanpa waktu gejala, padahal waktu gejala yang sebenarnya ialah lebih dari 24 jam setelah vaksin. Kemudian terdapat error dimana 9 orang yang diprediksi punya waktu gejala lebih dari 24 jam yang mana waktu gejala sebenarnya ke-9 orang tersebut ialah tanpa waktu gejala.



Gambar 9. Confusion Matrix pada model Random Forest

Dalam model *Random Forest* bisa kita lihat bahwa terdapat error dimana terdapat 14 orang yang diprediksi tanpa waktu gejala, padahal waktu gejala yang sebenarnya ialah lebih dari 24 jam setelah vaksin. Kemudian terdapat error dimana 9 orang yang diprediksi punya waktu gejala lebih dari 24 jam yang mana waktu gejala sebenarnya ke-9 orang tersebut ialah tanpa waktu gejala.



Gambar 10. Confusion Matrix pada model Support Vector Machine

Dalam model *Support Vector Machine* bisa kita lihat bahwa terdapat error dimana terdapat 15 orang yang diprediksi tanpa waktu gejala, padahal waktu gejala yang sebenarnya ialah lebih dari 24 jam setelah vaksin. Kemudian terdapat error dimana 12 orang yang diprediksi punya waktu gejala lebih dari 24 jam yang mana waktu gejala sebenarnya ke-9 orang tersebut ialah tanpa waktu gejala.

Berdasarkan dengan *precision, recall, f1-score dan accuracy* bisa dari masing-masing model bisa kita lihat:

Table 3. Evaluasi Model K-Nearest Neighbor

|  |  |  |  |
| --- | --- | --- | --- |
|  | *Precision* (%) | *Recall* (%) | F1-score (%) |
| 0 | 40 | 37 | 39 |
| 1 | 46 | 38 | 41 |
| 2 | 65 | 75 | 70 |
| *Accuracy (%)* | | | 55 |

Table 4. Evaluasi Model Random Forest

|  |  |  |  |
| --- | --- | --- | --- |
|  | *Precision* (%) | *Recall* (%) | F1-score (%) |
| 0 | 43 | 35 | 38 |
| 1 | 51 | 50 | 50 |
| 2 | 72 | 78 | 75 |
| *Accuracy (%)* | | | 60 |

Table 5. Evaluasi Model Support Vector Machine

|  |  |  |  |
| --- | --- | --- | --- |
|  | *Precision* (%) | *Recall* (%) | F1-score (%) |
| 0 | 41 | 44 | 42 |
| 1 | 50 | 38 | 43 |
| 2 | 70 | 79 | 74 |
| *Accuracy (%)* | | | 59 |

Berdasarkan tabel evaluasi ketiga model. Ketiga model memiliki *precision* dan *recall* hampir memiliki tinggi yang sama, saat memprediksi waktu gejala dengan value “2” atau gejala lebih dari 24 jam sehingga menghasilkan f1\_score yang lumayan tinggi terutama pada model *Random Forest*. Kemudian jika dibandingkan f1\_score dari ketiga model, saat melakukan prediksi pada responden yang memiliki waktu gejala kurang dari 24 jam (bernilai “1”) maka model *Random Forest* memiliki nilai f1\_score yang tinggi, lebih baik dibanding model *Support Vector Machine* di urutan ke-dua dan *K-Nearest Neighbor* di urutan ke-tiga. Kemudian saat dilakukan prediksi pada responden yang tidak memiliki gejala, model *Support Vector Machine* lebih baik dibanding kedua model lainnya karena memiliki f-1 score yang tinggi. Kemudian berdasarkan *accuracy*, model *Random Forest* lebih baik dibanding kedua model yang lainnya karena memiliki *accuracy* yang tinggi.

**KESIMPULAN**

Berdasarkan percobaan ke-tiga model uji yaitu model *K-Nearest Neighbor*, *Random Forest* dan *Support Vector Machine* pada kasus ini, model *Random Forest* lebih baik dibanding *Support Vector Machine* (di peringkat 2 dari terbaik) dan *K-Nearest Neighbor* (di peringkat 3 dari terbaik). Hal ini juga untuk mengurangi resiko agar pemilihan keputusan tidak membawa dampak yang terlalu merugikan staff rumah sakit. Dalam hal ini juga tidak terdapat perbedaan yang signifikan dari ke-tiga model karena perlakuan dataset yang “sama”. Saran untuk ke depan sebaiknya perlu dilakukan perlakuan data yang berbeda saat processing atau seleksi untuk setiap model hingga didapat perbedaan yang signifikan dan hasil yang lebih optimal.

**DAFTAR PUSTAKA**

Berrar, D. (2018). Cross-validation. *Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics*, *1*–*3*(January 2018), 542–545. https://doi.org/10.1016/B978-0-12-809633-8.20349-X

Bramasta, D. B. (2021). Bagaimana Upaya Pemerintah Yakinkan Masyarakat agar Mau Divaksin Covid-19? *Kompaspedia*. https://www.kompas.com/tren/read/2021/01/15/091200765/bagaimana-upaya-pemerintah-yakinkan-masyarakat-agar-mau-divaksin-covid-19-?page=all

COVID-19, S. T. P. (2021, November 18). Pemerintah Kejar Target 70% Populasi Indonesia Mendapat Vaksin. *Covid19.Go.Id*. https://covid19.go.id/p/berita/pemerintah-kejar-target-70-populasi-indonesia-mendapat-vaksin

Farisa, F. C. (2020). Alasan Jokowi Putuskan Vaksin Covid-19 Digratiskan untuk Masyarakat... *Nasional.Kompas.Com*. https://nasional.kompas.com/read/2020/12/17/09084331/alasan-jokowi-putuskan-vaksin-covid-19-digratiskan-untuk-masyarakat?page=all

Han, J., & Kamber, M. (2006). Data Mining: Concepts and Techniques. In A. Stephan (Ed.), *Morgan Kaufmann* (Second, Vol. 12). Diane Cerra. https://doi.org/10.1007/978-3-642-19721-5

Handayani, R. T., Kuntari, S., Darmayanti, A. T., Widiyanto, A., & Atmojo, J. T. (2020). Factors Causing Stress in Health and Community When the Covid-19 Pandemic. *Jurnal Keperawatan Jiwa*, *8*(3), 353. https://doi.org/10.26714/jkj.8.3.2020.353-360

Kementerian Kesehatan. (2021). Question ( Faq ) Pelaksanaan Vaksinasi Covid-. *2020*, *2*(1), 1–16. https://kesmas.kemkes.go.id/assets/uploads/contents/others/FAQ\_VAKSINASI\_COVID\_\_call\_center.pdf

Permatasari, D. (2021). Kebijakan Covid-19 dari PSBB hingga PPKM Empat Level. *Kompaspedia*, 1–18. https://kompaspedia.kompas.id/baca/infografik/kronologi/kebijakan-covid-19-dari-psbb-hingga-ppkm-empat-level

Rachman, F., Health, S. P.-I. of, & 2020, undefined. (2020). Analisis Sentimen Pro dan Kontra Masyarakat Indonesia tentang Vaksin COVID-19 pada Media Sosial Twitter. *Inohim.Esaunggul.Ac.Id*, *8*(2), 2655–9129. https://inohim.esaunggul.ac.id/index.php/INO/article/download/223/175

Ramadhan, A. (2021). *Menkes Ungkap Alasan Tenaga Kesehatan Jadi Prioritas Vaksinasi Covid-19*. https://nasional.kompas.com/read/2021/01/11/09112231/menkes-ungkap-alasan-tenaga-kesehatan-jadi-prioritas-vaksinasi-covid-19

Rokom. (2021a). *Positif COVID-19 Uasi Divaksinasi? Begini Penjelasan Komnas KIPI dan Kemenkes*. https://sehatnegeriku.kemkes.go.id/baca/umum/20210222/1237036/\_\_trashed-2/

Rokom. (2021b, November 5). *Vaksinasi COVID-19 di Indonesia Capai 200 Juta Suntikan*. https://sehatnegeriku.kemkes.go.id/baca/umum/20211105/1038788/vaksinasi-covid-19-di-indonesia-capai-200-juta-suntikan/

Rokom. (2021c, November 17). Survei Tunjukkan Mayoritas Masyarakat Indonesia Bersedia Menerima Vaksin COVID-19. *Sehatnegeriku.Kemkes.Go.Id*. https://sehatnegeriku.kemkes.go.id/baca/umum/20201117/4935712/survei-tunjukkan-mayoritas-masyarakat-indonesia-bersedia-menerima-vaksin-covid-19-2/

Shereen, M. A., Khan, S., Kazmi, A., Bashir, N., & Siddique, R. (2020). COVID-19 infection: Origin, transmission, and characteristics of human coronaviruses. *Journal of Advanced Research*, *24*, 91–98. https://doi.org/10.1016/j.jare.2020.03.005

Singh, H. (2019). *Understanding Random Forests*. https://medium.com/@harshdeepsingh\_35448/understanding-random-forests-aa0ccecdbbbb

Speiser, J. L., Miller, M. E., Tooze, J., & Ip, E. (2019). A comparison of *Random Forest* variable selection methods for classification prediction modeling. *Expert Systems with Applications*, *134*, 93–101. https://doi.org/10.1016/j.eswa.2019.05.028

Urutan Daftar Prioritas Penerima Vaksin Covid-19. (2020). *CNNIndonesia*. https://www.cnnindonesia.com/nasional/20201229102016-20-587293/urutan-daftar-prioritas-penerima-vaksin-covid-19