

File1

by Hap1 Hap2

Submission date: 03-Apr-2020 10:26AM (UTC+0700)

Submission ID: 1288383747

File name: blind_version_submitKEMAS.docx (1.47M)

Word count: 3251

Character count: 20081

Menelusuri Covid19 di Dunia dan di Indonesia dengan Model Regresi SVM, Bayesian dan Gaussian

Abstract. Pada artikel ini ditunjukkan penelusuran data total kasus Covid19 di dunia dan di Indonesia. Metode yang digunakan adalah metode SVM (*Support Vector Machine*) dan regresi Bayesian Ridge dimana pengaruh *Social Distancing* belum ditunjukkan sehingga digunakan model eksponensial dan Gaussian. Metode SVM dan regresi Bayesian Ridge berhasil memprediksi 10 hari selanjutnya dimana pendekatan yang terbaik diberikan oleh Bayesian Ridge. Pada model Gaussian diperoleh prediksi dalam bentuk 2 skenario yaitu skenario terburuk dan skenario realistis. Skenario terburuk memberikan total kasus pada 24 April 2020 sebesar sekitar 14675. Hal ini nampaknya tidak realistis jika diperhatikan data terakhir pada 31 Maret 2020 sebesar 1677. Oleh karena itu diusulkan skenario realistis dimana total kasus pada puncak berlaku pada 24 April 2020 dengan total kasus 4242.

Kata Kunci : Covid19, SVM, Bayesian Ridge , Gaussian , Social Distancing

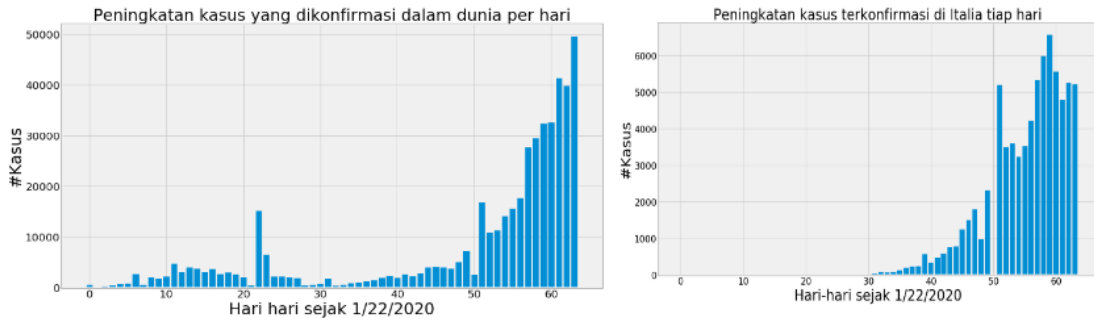
LATAR BELAKANG

Telah diketahui bahwa munculnya virus Covid-19 diumumkan pertama kali di Wuhan pada Desember 2019. Laju penularan yang sangat cepat oleh karena mobilitas masyarakat dunia yang tinggi menyebabkan kesiapan negara-negara di dunia dipandang tidak cukup dalam mengantisipasi penularan yang terjadi. Adanya pendataan yang cepat dari John Hopkins University membantu untuk mencermati penyebaran virus ini pada berbagai negara di dunia. Beberapa lembaga statistik di dunia juga melakukan pendataan yang selalu diperbaharui, misalkan (Ortiz-Ospina, 2020). Demikian pula kepedulian sosial muncul atas pengaruh yang ditimbulkan akibat perkembangan laju Covid19 yang meningkat sangat cepat (Hamid, 2020). Beberapa pendataan kemudian diilustrasikan pada grafik sehingga dapat dicermati sebagaimana ditunjukkan pada Gambar 1 dimana WHO mencatat kejadian Covid19 sejak 22 Januari 2020 hingga 16 Maret 2020 di berbagai belahan dunia. Salah satu yang terlihat peningkatan kasus yang sangat tinggi selain Wuhan di China juga terjadi di Italia dan US sebagaimana pada Gambar 1(kanan)-2(kiri). Selanjutnya kita juga dapat mengilustrasikan perbandingan total konfirmasi ketiga negara tersebut pada Gambar 2(kanan). Demikian pula diperoleh data perkembangan kasus Covid19 di 10 negara terbanyak yang ditunjukkan pada Gambar 3.

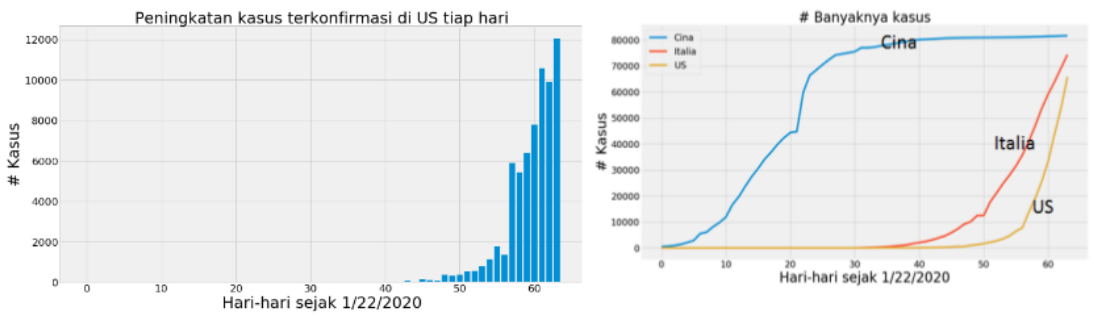
Pendataan secara cepat dan *uptodate* ini telah dilakukan dan diharapkan dapat memberikan penjelasan kepada semua pihak untuk dapat segera melakukan tindakan pencegahan dimana sekarang diberlakukan *Social Distancing* (SD) ataupun *Physical Distancing* (PD) yaitu pembatasan interaksi sosial antar manusia maupun kegiatan yang melibatkan kelompok. Sejak 16 Maret 2020 di Indonesia SD atau PD diberlakukan. Telah direkomendasikan oleh penelitian bagaimana SD ataupun PD berpengaruh terhadap memperlambat laju kasus Covid19 di Wuhan (Prem et al., 2020) serta di seluruh dunia melakukan SD (Pambuccian, 2020) hingga 2 April 2020 (hingga artikel ditulis dalam situasi SD). Aksi SD ini telah dilakukan dan laju pertumbuhan pasien yang terkonfirmasi Covid-19 perlu dipelajari untuk melihat ada tidaknya pengaruh yang signifikan baik pada pendataan di dunia maupun di Indonesia. Untuk itu, pada tulisan ini ditunjukkan penelusuran pendataan kasus yang dikonfirmasi serta yang meninggal pada tingkat dunia dimana data di WHO di mulai pada 22 Januari 2020 sedangkan data di Indonesia dimulai dari 6 Maret 2020. Metode yang digunakan adalah SVM (*Support Vector Machine*) dan regresi Bayesian Ridge. Metode tersebut biasa digunakan dalam melakukan klasifikasi yang biasa dilakukan dengan Python (Min-Wei et al., 2018)(Lamidi, 2018)(Mckinney, 2018) sebagai metode yang banyak digunakan pada Machine Learning pada pengolahan big data (Demidova et al., 2016).

Akan tetapi masyarakat menghendaki bahwa model yang digunakan dapat menunjukkan waktu-waktu dimana terjadi kasus Covid19 mencapai puncak dan kapan terjadi kasus Covid19 menurun karena adanya pengaruh SD. Untuk itulah model eksponensial juga diperkenalkan yang kemudian digunakan sebagai bagian dari model Gaussian. Hal

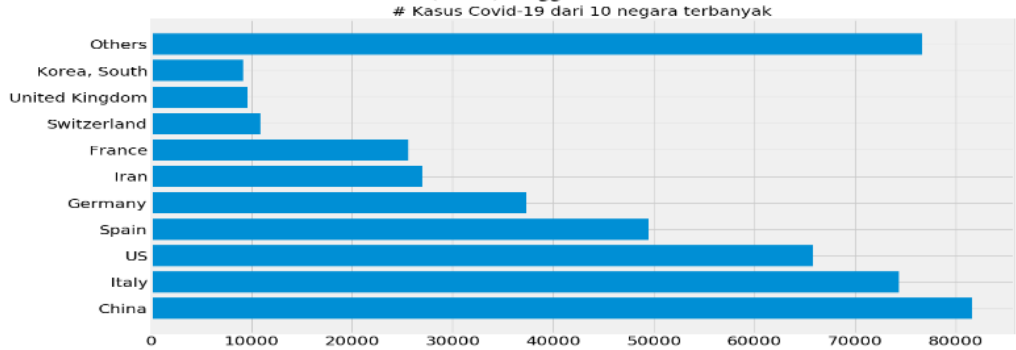
inilah yang menjadi novelty atas penelitian ini dimana model dapat memberikan prediksi banyaknya total kasus yang muncul pada waktu tertentu.



GAMBAR 1. Laju peningkatan kasus yang dikonfirmasi Banyaknya kasus (sumbu vertikal), banyaknya hari (sumbu horizontal) di dunia (kiri) dan di Italia yang dicatat WHO sejak 22 Januari 2020 hingga 16 Maret 2020 .



GAMBAR 2. Laju peningkatan kasus yang dikonfirmasi Banyaknya kasus terkonfirmasi (sumbu vertikal), banyaknya hari (sumbu horizontal) di dunia yang dicatat WHO sejak 22 Januari 2020 di US (kiri) dan perbandingan data dari 3 negara yaitu Cina, Italia, US (kanan) hingga 16 Maret 2020.



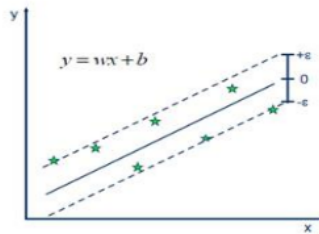
GAMBAR 3. Banyak kasus terkonfirmasi di 10 negara terbanyak (sumbu vertikal), banyaknya hari (sumbu horizontal) yang dicatat WHO sejak 22 Januari 2020 hingga 16 Maret 2020.

METODE PENELITIAN

Model Regresi SVM

1

SVM merupakan salah satu metode klasifikasi dalam *machine learning*. Tetapi kemudian SVM juga dapat digunakan untuk regresi yang disebut juga Akan tetapi dapat pula *Support Vector Regression* (SVR) mencocokkan sebanyak mungkin data dengan garis serta memperhatikan margin sebesar epsilon sebagaimana ditunjukkan pada Gambar 4.



GAMBAR 4. Ilustrasi SVM dengan margin selebar ϵ

1 SVR menggunakan prinsip yang sama dengan SVM untuk klasifikasi dengan hanya sedikit perbedaan. Ide dasar SVR sama dengan SVM yaitu : meminimalkan error, pembuatan *hyperplane* yang memaksimalkan margin. Masalah optimasi pada SVM dengan kernel ϕ yang memetakan x_i ke $\phi(x_i)$ adalah

$$\min_{w, b, \zeta} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \zeta_i \quad (1)$$

dengan kendala

$$y_i(w \cdot \phi(x_i) + b) \geq 1 - \zeta_i \text{ untuk semua } 1 \leq i \leq n, \zeta_i \geq 0. \quad (2)$$

Kernel ini merupakan fungsi yang dapat dipilih seperti polinomial, sigmoid yang disediakan dalam paket python untuk mendefinisikan kurva pemisah data. Demikian pula parameter C pada fungsi tujuan merupakan input pada fungsi yang sudah ada.

Model Regresi Bayesian

Model linear regresi Bayesian Ridge merupakan perpaduan antara regresi Ridge dan linear regresi Bayesian (Permai & Tanty, 2018) (Qian, 2018). Regresi Ridge menggunakan regularisasi dengan norm L_2 sedangkan regresi Bayesian model didefinisikan secara probabilistik dengan prior eksplisit untuk parameter-parameter. Pemilihan prior berpengaruh dalam regularisasi. Jadi keduanya tidak sama karena ridge regresi adalah suatu model regresi, dan pendekatan Bayesian adalah langkah yang lebih umum untuk mendefinisikan dan mengestimasi model statistik yang dapat diaplikasikan pada model-model yang berbeda.

Diketahui bahwa regresi standar dalam 1 variabel dapat ditulis sebagai $y = \beta_0 + \beta_1 x + \epsilon$

Untuk mengubah regresi standar ini dalam regresi Bayesian maka regresi standar tersebut ditulis dalam model probabilistik dalam bentuk

$$\mu_i = \alpha + \beta x_i$$

$$y_i \sim \mathcal{N}(\mu_i, \sigma)$$

Yang berarti variabel dependen Y mengikuti distribusi normal dengan parameter μ_i yaitu fungsi linear X yang diparameterisasi oleh α, β dan deviasi standard σ . Jika kita menggunakan model linear dengan OLS (*Ordinary Least Square*) kita tidak perlu memperlakukan tentang probabilitas karena kita mencari nilai optimal α, β dengan meminimalkan kuadrat error dalam pencocokan data dengan pendekatan. Sebaliknya, kita dapat menduga model tersebut dengan memaksimalkan fungsi *likelihood*

$$\arg \max_{\alpha, \beta, \sigma} \prod_{i=1}^n \mathcal{N}(y_i; \alpha + \beta x_i, \sigma)$$

dimana \mathcal{N} adalah fungsi densitas dari distribusi normal yang dihitung pada titik-titik y_i dengan parameter $\alpha + \beta x_i$ dan σ .

Pada pendekatan Bayesian, sebagai pengganti memaksimalkan fungsi likelihood, kita akan mengasumsikan distribusi prior untuk parameter dan menggunakan teorema Bayesian

$$\text{posterior} \propto \text{likelihood} \times \text{prior}$$

Fungsi likelihood sama dengan di atas, tetapi yang berubah adalah bahwa kita mengasumsikan bahwa distribusi prior untuk parameter-parameter α, β, σ yang diduga dan melibatkannya dalam persamaan

$$f(\alpha, \beta, \sigma | Y, X) \propto \prod_{i=1}^n \mathcal{N}(y_i; \alpha + \beta x_i, \sigma) f_\alpha(\alpha) f_\beta(\beta) f_\sigma(\sigma).$$

Distribusi yang dipilih banyak untuk α, β .

Penyusunan MLE dan MAP

Terdapat 2 error yang digunakan yaitu MLE dan MAP. MLE (*Maximum Likelihood Estimation*) dan MAP (*Maximum a posterior probability*) dimana MLE memaksimalkan keserupaan data yang ditentukan parameter dan MAP memilih parameter yang memaksimalkan probabilitas posterior berdasarkan data.

Sebutlah himpunan pasangan data $D = \{x_i, y_i\}$ dimana $x_i \in \mathcal{R}^n$ dan $y_i \in \mathcal{R}$. Model linear regresi menduga nilai y_i sebagai kombinasi linear dari variabel x_i dimana

$$y_i = w_0 + \sum_j w_j x_{ij} = w_0 + \mathbf{w}^T \mathbf{x}_i \quad (*)$$

Dengan menambahkan 1 sebagai vektor pada x_i dan menggabungkan w_0 dan w_i pada satu vektor, persamaan (*) dapat ditulis dalam bentuk

$$y_i = \mathbf{w}^T \mathbf{x}_i$$

Nilai observasi y_i diasumsikan mempunyai derau Gaussian, sehingga ditulis

$$y_i = \mathbf{w}^T \mathbf{x}_i + \epsilon \text{ dimana } \epsilon \sim \mathcal{N}(0, \sigma^2)$$

Fungsi *likelihood* untuk kasus ini adalah

$$\mathcal{L}(D | \mathbf{w}, \sigma) = (2\pi\sigma^2)^{-n/2} \prod_i \exp\left[-\frac{(y_i - \hat{y}_i)^2}{2\sigma^2}\right]$$

Sehingga logaritma L adalah

$$\ln \mathcal{L} = -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_i (y_i - \mathbf{w}^T \mathbf{x}_i)^2$$

Dengan MLE berarti \mathbf{w} diberikan oleh

$$\mathbf{w}_{MLE} = \arg \max_w \ln \mathcal{L}(D|w, \sigma)$$

yang menjadi

$$\mathbf{w}_{MLE} = \arg \min_w \sum_i (y_i - w^T x_i)^2$$

Dengan menurunkan terhadap \mathbf{w} dan membuat sama dengan 0 maka kita dapat memperoleh

$$\mathbf{w} = (X^T X)^{-1} X^T \mathbf{y}$$

Dengan MAP, kita mengasumsikan prior Gaussian untuk \mathbf{w} yaitu $w \sim \mathcal{N}(0, \lambda^{-1}I)$

$$P(w) = \left(\frac{\lambda}{2\pi}\right)^{n/2} \exp\left[-\frac{\lambda}{2} w^T w\right]$$

Jadi probabilitas posterior dengan aturan Bayes ditulis sebagai

$$P(w|D) = \frac{P(w)P(D|w)}{P(D)}$$

Sebagai MLE, kita akan memaksimalkan logaritma probabilitas posterior dan

$$\mathbf{w}_{MAP} = \arg \max_w \ln P(w|D)$$

yang menjadi

$$\mathbf{w}_{MAP} = \arg \min_w \sum_i (y_i - w^T x_i)^2 + \frac{\lambda}{2} w^T w.$$

Jadi dengan estimasi MAP dapat sebagai suatu regresi dengan regularisasi. Dugaan MAP pada kasus ini adalah

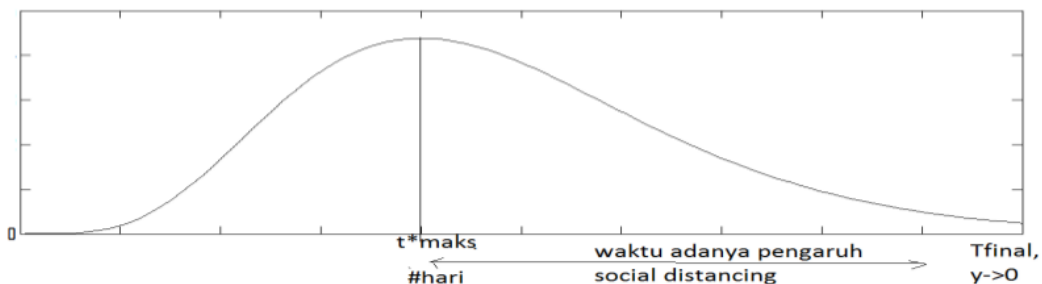
$$\mathbf{w}_{MAP} = (\lambda I + X^T X)^{-1} X \mathbf{y}.$$

Model Eksponensial dan Gaussian

Terdapat beberapa model eksponensial yang digunakan tetapi kita dapat memilih bentuk umum yang menunjukkan bahwa total kasus Covid19 dapat diperoleh, yaitu

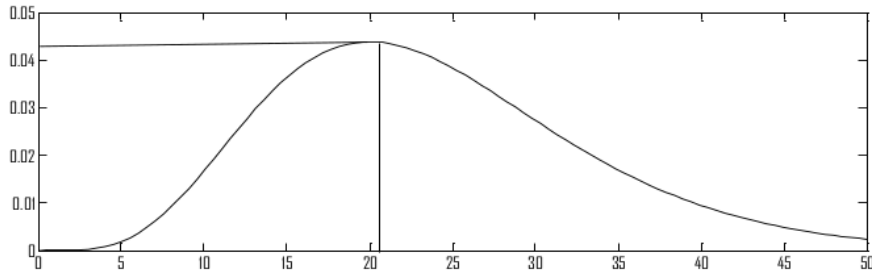
$$y(t) = a_1 + a_2 \exp(a_3 t) \quad (3)$$

Jika model eksponensial digunakan, maka pada waktu tak hingga total kasus Covid19 juga akan positif tak hingga dimana hal itu tidak mungkin karena upaya pemerintah dalam gerakan *Social Distancing* tidak memberi dampak sama sekali. Untuk itu model perlu diperbaiki. Kita berharap bahwa *Social Distancing* atau sekarang lebih disebut dengan *physical distancing* mempunyai pengaruh pada penurunan kasus sebagaimana ditunjukkan pada Gambar 5 dimana model menyerupai Gaussian digunakan.



GAMBAR 5. Ilustrasi model fungsi Covid 19 di Indonesia sebagai fungsi waktu.

Hingga saat ini kita belum tahu t^* maks yaitu waktu dimana total kasus kumulatif di Indonesia mengalami penurunan . Tetapi kita berharap model Gambar 5 memberikan ilustrasi yang berlaku untuk semua kasus Covid19 yang berhasil diturunkan pandemiknya. Misalkan sekarang bahwa untuk data tak berdimensi, Gambar 5 diperoleh dengan skala pada sumbu horisontal dan vertikal yang ditunjukkan pada Gambar 6.



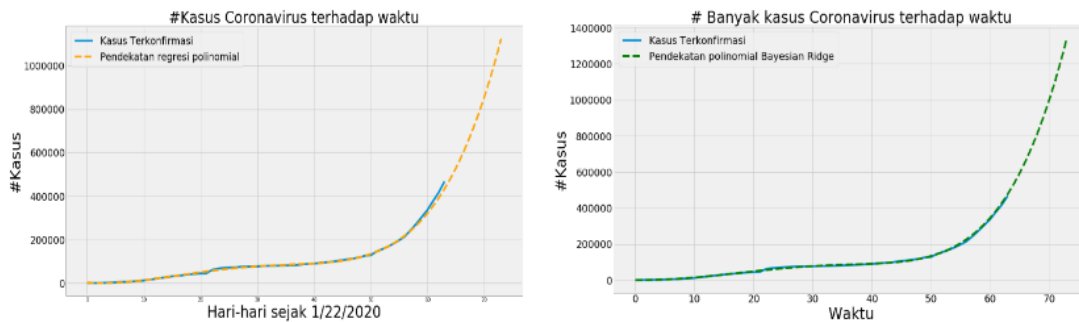
GAMBAR 6. Ilustrasi model fungsi Covid 19 yang akan dicocokkan skalanya dengan waktu yang ada.

Perhatikan Gambar 5 dan Gambar 6. Kita ingin menduga dimanakah fungsi eksponensial berada pada Gambar 5. jika fungsi eksponensial diterapkan dan menempati pada Gambar 6. Hal inilah yang dilakukan dalam metode ini. Pada model ini data yang digunakan b adalah data pada 6 Maret-27 Maret 2020 dan 6 Maret-31 Maret 2020 .

HASIL DAN PEMBAHASAN

Hasil SVM dan Ridge Bayesian untuk banyaknya total kasus di dunia

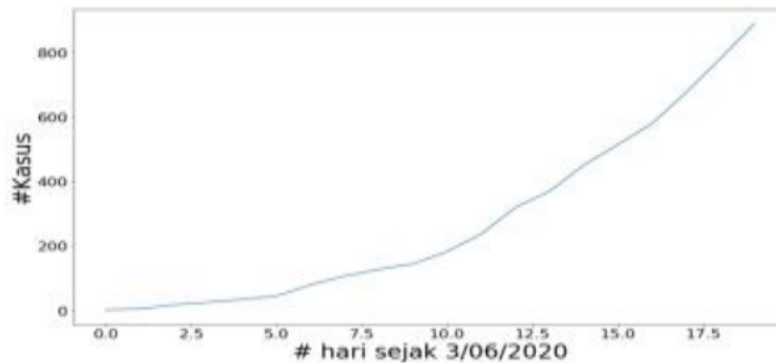
Dengan menggunakan formulasi (1)-(2) dan bantuan Python kita dapat melakukan pendekatan dan prediksi banyaknya total kasus di dunia pada 22 Januari 2020 hingga 16 Maret 2020. Kemudian dilakukan dugaan untuk 10 hari selanjutnya. Kedua metode SVM dan Bayesian Ridge memberikan prediksi yang menyerupai dimana semua menunjukkan peningkatan secara eksponensial. Akan tetapi MLE dan MAP untuk Bayesian Ridge lebih kecil dimana pada Gambar 4 terlihat bahwa metode Bayesian Ridge lebih baik.



GAMBAR 4. Laju peningkatan kasus yang dikonfirmasi
Banyaknya kasus terkonfirmasi (sumbu vertikal), banyaknya hari (sumbu horisontal)
di dunia yang dicatat WHO sejak 22 Januari 2020 dibandingkan dengan regresi polinomial (kiri) dan Bayesian Ridge (kanan) .

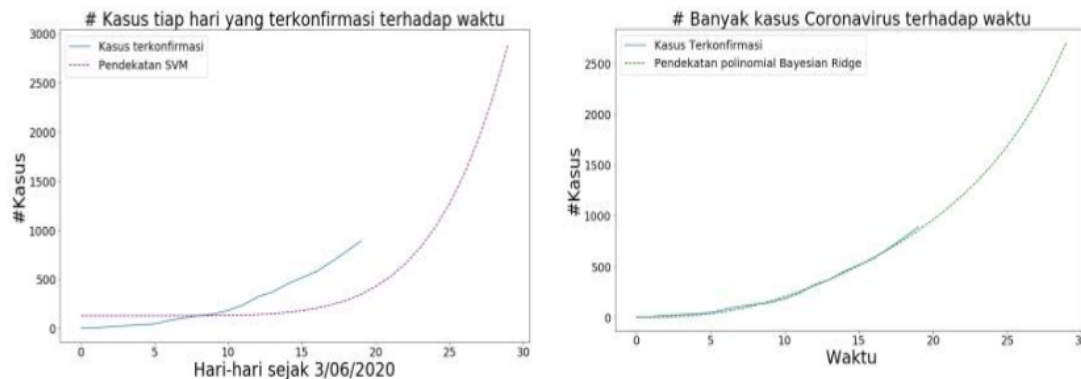
Hasil SVM dan Ridge Bayesian untuk banyaknya total kasus di Indonesia

Metode SVM dan Bayesian Ridge telah diujicobakan untuk memodelkan data *timeseries* untuk sebagian besar negara-negara yang terkonfirmasi adanya paparan virus Covid-19 dengan data dari 22 Januari 2020 hingga 16 maret 2020 yang ditunjukkan pada Gambar 5.



GAMBAR 5. Banyaknya kasus (sumbu vertikal), banyaknya hari (sumbu horizontal) di Indonesia yang dicatat CNN /dilaporkan Metro TV sejak 6 maret 2020-26 Maret 2020 .

Selanjutnya dengan metode SVM dan Bayesian Ridge dipelajari hingga 10 hari berikutnya. Telah ditunjukkan bahwa Bayesian Ridge memberikan pendekatan terbaik untuk berbagai data dari seluruh dunia maupun dari 3 negara terbanyak yang terpapar Covid-19. Dengan ide pendekatan yang sama, maka data dari Indonesia sejak 6 maret 2020 hingga 26-27 maret 2020 dianalisa. Hasil pendekatan Bayesian Ridge memberikan simulasi yang lebih mendekati data pada Gambar 6 dibandingkan pendekatan SVM. Oleh karena pada kasus dunia dan Indonesia, metode Bayesian Ridge lebih baik, maka data pasien yang meninggal didekati juga menggunakan Bayesian Ridge yang ditunjukkan pada Gambar 7. Akan tetapi dua metode menunjukkan masih adanya peningkatan 10 hari ke depan setelah 26 maret 2020 dengan pendekatan kedua metode.



GAMBAR 6(kiri). Banyaknya kasus yang dikonfirmasi (sumbu vertikal), banyaknya hari (sumbu horizontal) di Indonesia yang dicatat CNN / dilaporkan Metro TV sejak 6 maret 2020 hingga 26 Maret 2020 dan selanjutnya dibandingkan dengan SVM .

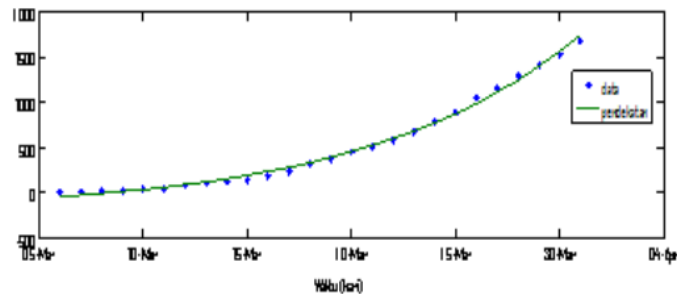
(kanan). Pendekatan kasus yang dikonfirmasi (sumbu vertikal), banyaknya hari (sumbu horizontal) di Indonesia yang dicatat CNN sejak 6 maret 2020 / dilaporkan Metro TV dengan Bayesian Ridge dibandingkan dengan data



GAMBAR 7. Pendekatan pasien Covid19 yang meninggal di Indonesia sejak 6 Maret 2020-27 Maret 2020 (sumbu vertikal), banyaknya hari (sumbu horizontal) di Indonesia yang dicatat CNN dan berita metro TV sejak 6 maret 2020-27 maret 2020 dengan Bayesian Ridge dibandingkan dengan data

Hasil Model Eksponensial

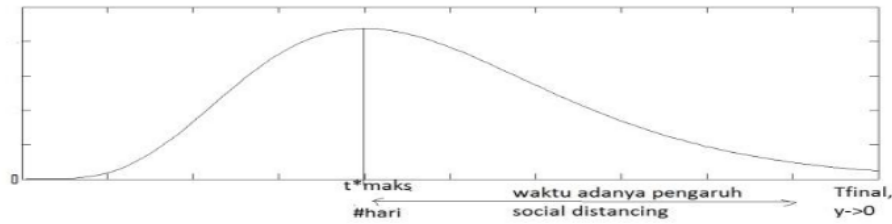
Beberapa model pertumbuhan Covid19 di Indonesia telah dilakukan oleh beberapa peneliti di Indonesia maupun komunitas yang terkait dimana laju pertumbuhan telah mengikuti model eksponensial yang beredar di internet. Salah satu model yang bisa digunakan fungsi eksponensial pada persamaan(3). Dengan melakukan fitting terhadap data yang ada kita dapat memperoleh pula hasil yang cocok dengan data dengan error fungsi 4.38 % yang ditunjukkan pada Gambar 8. Pendekatan ditunjukkan pada data dengan data dari 6 Maret-31 Maret 2020. Akan tetapi dengan model ini, pengguna masih memerlukan waktu selanjutnya beserta prediksi mendatang bagaimana pengaruh *Social Distancing* terhadap penurunan total kasus Covid19 di Indonesia. Untuk itulah model eksponensial pada Gambar 8 tidak lagi relevan sekalipun model cukup bagus berdasarkan data yang ada padahal data pada mendatang belum ada. Prediksi demikian dibutuhkan dalam melakukan antisipasi serta menyusun berbagai kebijakan yang dilakukan pemerintah. Data menunjukkan bahwa laju penambahan kasus perhari masih meningkat demikian pula laju total kasus sebagaimana ditunjukkan pada berita metro TV bahwa total kasus 1790 pada 2 April 2020 .



GAMBAR 8. Model pendekatan eksponensial untuk data total konfirmasi data per hari di Indonesia dari 6 maret 2020-31 maret 2020.

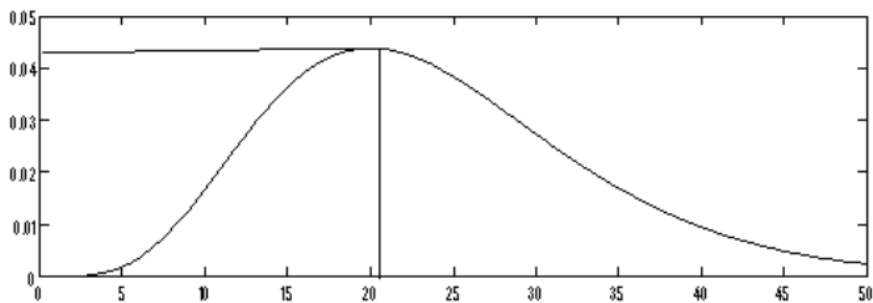
Hasil model Gaussian beserta prediksi waktu

Kita berharap bahwa *Social Distancing* atau sekarang lebih disebut dengan *Physical Distancing* mempunyai pengaruh pada pada penurunan kasus sebagaimana ditunjukkan pada Gambar 9 dimana model menyerupai Gaussian digunakan.



GAMBAR 9 Ilustrasi model fungsi Covid 19 di Indonesia sebagai fungsi waktu.

Hingga saat ini kita belum tahu t^*_{maks} yaitu waktu dimana total kasus kumulatif di Indonesia mengalami penurunan. Tetapi kita berharap model Gambar 9 memberikan ilustrasi yang berlaku untuk semua kasus Covid19 yang berhasil diturunkan pandemiknya. Misalkan sekarang bahwa untuk data tak berdimensi, Gambar 9 diperoleh dengan skala pada sumbu horisontal dan vertikal dengan Gambar 10.

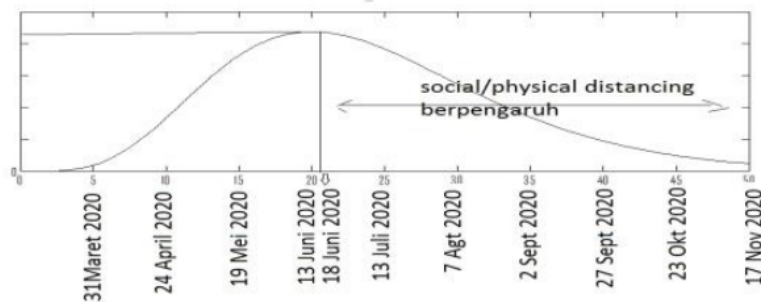


GAMBAR 10. Ilustrasi model fungsi Covid 19 yang akan dicocokkan skalanya dengan waktu yang ada.

Perhatikan Gambar 8 dan Gambar 10. Kita ingin menduga dimanakah fungsi eksponensial berada pada Gambar 10. jika Gambar 8 diskalakan pada Gambar 10. Terdapat 2 hasil prediksi yang diperoleh yang dijelaskan pada berikut ini.

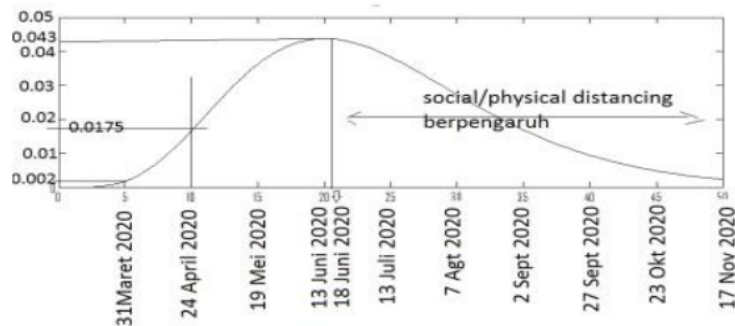
Skenario Terburuk Waktu Pertumbuhan Covid 19 Di Indonesia

Jika dalam 25 hari berlaku data awal Indonesia (sebenarnya digunakan 26 hari, untuk mempermudah penskalaan), pada skala 0-5 pada Gambar 10 maka pada point sumbu horisontal adalah 5 , maka waktu terakhir pada Gambar 8 berada pada tanggal 31 Maret 2020. Karena setiap penambahan 5 satuan pada Gambar 8 berarti 25 hari maka pada puncak berlaku sekitar 105 hari. Demikian dilakukan maka diperoleh skenario kasus Covid19 di Indonesia dengan waktunya ditunjukkan pada Gambar 11.



GAMBAR 11. Ilustrasi skenario terburuk model fungsi Covid 19 di Indonesia yang akan dicocokkan skalanya dengan waktu yang ada dengan pengaruh social distancing sudah berpengaruh .

Dengan model itu maka kita dapat menentukan total kasus pada tanggal - tanggal tertentu di atas. Karena berdasarkan data diketahui bahwa pada 31 maret 2020 sebanyak 1677 kasus, maka kita sekarang dapat menduga pada tanggal yang lain. Misalkan pada 31 maret maka sumbu y bernilai 0.002 . Oleh karena itu pada 24 April 2020 akan berada pada sumbu y yang bernilai sekitar 0.0175. Dengan skala tersebut kita dapat memperkirakan bahwa banyaknya total kasus hingga 24 April 2020 sekitar 14675 kasus yang dihasilkan pada Gambar 12.

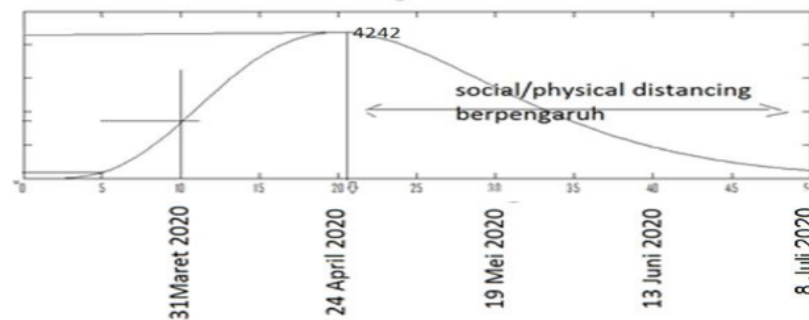


GAMBAR 12. Ilustrasi skenario model fungsi Covid 19 di Indonesia yang akan dicocokkan skalanya dengan waktu yang ada dengan pengaruh social distancing sudah berpengaruh dengan skala banyaknya total kasus

Nampaknya skenario terburuk tidak terjadi karena dalam 25 hari selanjutnya menjadi 14675 dianggap terlalu besar dari pengamatan data yang ada. Untuk itu kita bisa menggeser tanggal di sumbu horizontal dengan model berikut.

Skenario Kedua yang lebih realistis

Data dari 6 maret 2020-hingga 31 Maret 2020 berada pada skala 10 satuan pada sumbu horizontal. Dengan skenario ini maka diperoleh urutan tanggal dan puncak terjadi pada 24 April 2020 dan ditunjukkan pada Gambar 13.



GAMBAR 13. Ilustrasi skenario realistis dengan total kasus pada puncak di 24 April 2020 dengan adanya pengaruh *social distancing*

Dengan skenario lebih realistis ini maka kita dapat menduga banyaknya total kasus hingga pada 24 April 2020 sebesar 4242 dimana ini lebih cukup realistis setelah 25 hari dari 31 Maret 2020 dan ini merupakan waktu dimana terjadi puncak total kasus sehingga *Sosial Distancing* baru mulai memberikan dampaknya bagi pertumbuhan total kasus Covid 19 di Indonesia. Dibandingkan dengan penulisan dalam beberapa peneliti dalam media massa ((Ardila, n.d.) maka hasil ini sangat jauh karena pada April 2020 diprediks³ oleh peneliti lain dapat sekitar 500 ribu. Sedangkan menurut peneliti dari ITB menyatakan bahwa puncak terjadi pada akhir Maret 2020 dan berakhir pada pertengahan

April dengan kasus harian sekitar 600 (Adi, n.d.). Hingga pada 2 April 2020, kasus harian adalah 113 sehingga dapat dimungkinkan bahwa prediksi pada pertengahan April 2020 masih sangat jauh dari 600.

KESIMPULAN

Pada tulisan ini ditunjukkan penelusuran data total kasus Covid19 di dunia dan di Indonesia. Untuk data total kasus di dunia digunakan data sejak 22 Januari 2020 hingga 16 Maret 2020. Kemudian dilakukan pendekatan untuk mendapatkan total kasus 10 hari selanjutnya menggunakan SVM dan Bayesian Ridge. Kedua metode menunjukkan bahwa terjadi peningkatan pada 10 hari selanjutnya secara eksponensial dimana pendekatan Bayesian Ridge yang menunjukkan error MLE dan MAP lebih kecil. Selanjutnya kedua metode digunakan pada data total kasus Covid19 dari tanggal 6 maret 2020-27Maret 2020. Dengan kedua metode ini pengaruh *Social Distancing* belum dapat ditunjukkan. Untuk itulah digunakan model eksponensial dan Gaussian digunakan data pada tanggal maret 2020-31 Maret 2020. Model eksponensial digunakan sebagai inisialisasi bagi data model Gaussian yang memperhatikan *Social Distancing*. Model Gaussian *like* yang dihasilkan dengan skala yang sudah ada dan ditransformasikan sesuai dengan data. Model ini diusulkan dengan 2 skenario dimana skenario terburuk memberikan hasil bahwa *Social Distancing* mulai memberi dampak pada mulai sekitar 18 Juni 2020. Akan tetapi total kasus yang diprediksi pada 24 April 2020 sekitar 14675 kasus. Jika diperhatikan dari data yang sudah ada bahwa dalam 25 hari pada 6 maret-31 maret 2020, total kasus adalah 1677 maka prediksi skenario terburuk tidak dapat diterima. Untuk itu maka diusulkan skenario kedua dimana puncak total kasus Covid19 di Indonesia terjadi pada 24 April 2020 dengan total kasus 4242 dan itu terjadi dan sejak itu pula *Social Distancing* memberikan dampak.

4

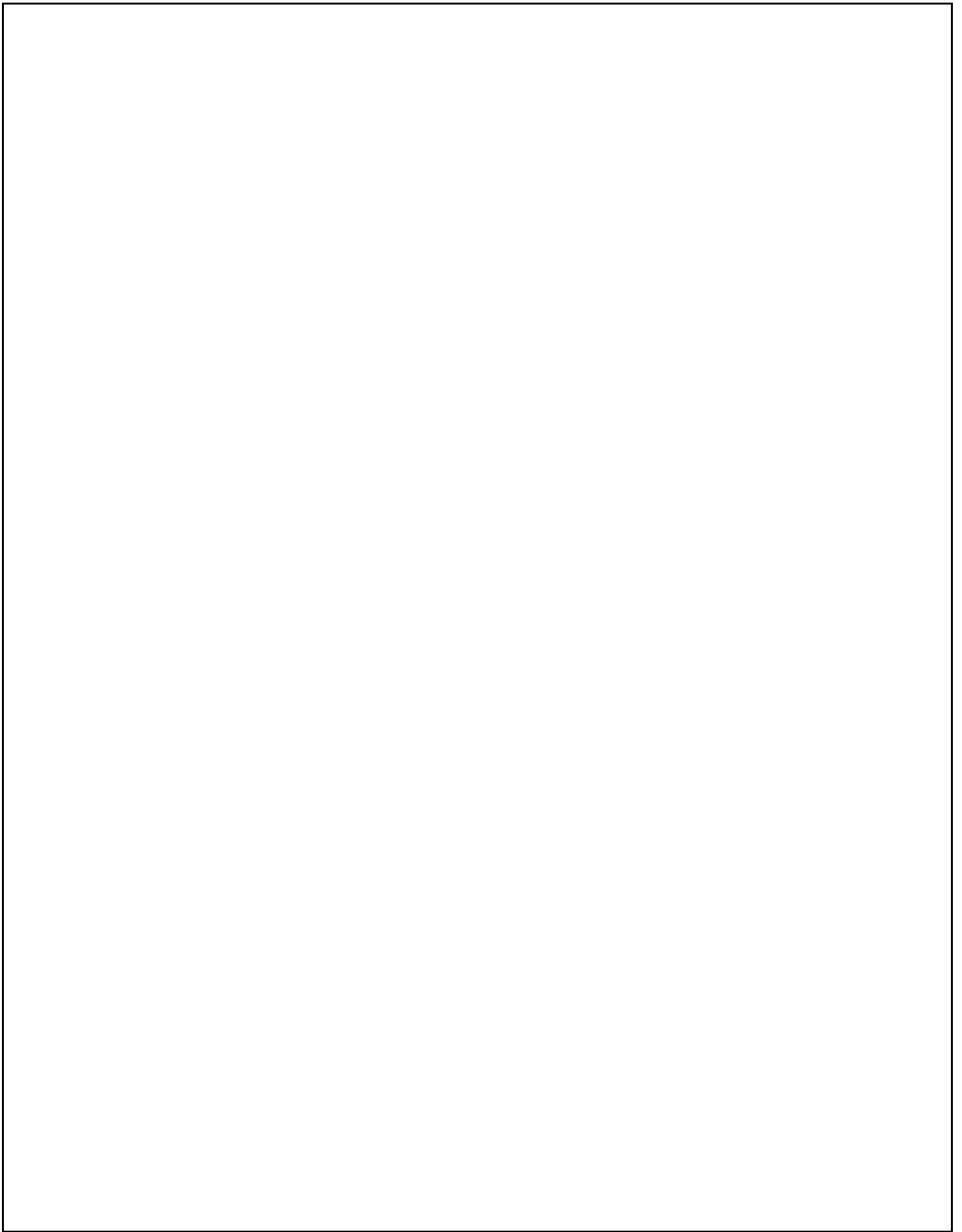
UCAPAN TERIMA KASIH

Penelitian ini didanai oleh pusat riset penelitian UKSW tahun anggaran 2019/2020 dalam rangka kepedulian terhadap kasus Covid19 di dunia dan di Indonesia.

REFERENCES

- Adi, P. (n.d.). *Peneliti ITB Buat Pemodelan Prediksi Puncak Penyebaran COVID-19 di Indonesia*. <https://www.itb.ac.id/news/read/57444/home/peneliti-itb-buat-pemodelan-prediksi-puncak-penyebaran-covid-19-di-indonesia>
- Ardila, S. (n.d.). *Bappenas, UI modeling shows grim projection of COVID-19 spread in Indonesia*. <https://www.thejakartapost.com/news/2020/03/31/bappenas-ui-modeling-shows-grim-projection-of-covid-19-spread-in-indonesia.html>
- Demidova, L., Nikulchev, E., & Sokolova, Y. (2016). Big Data Classification Using the SVM Classifiers with the Modified Particle Swarm Optimization and the SVM Ensembles. *International Journal of Advanced Computer Science and Applications*, 7(5), 294–312. <https://doi.org/10.14569/ijacsa.2016.070541>
- Hamid, A. R. A. H. (2020). Social responsibility of medical journal: a concern for COVID-19 pandemic. *Medical Journal of Indonesia*, 29(1), 1–3. <https://doi.org/10.13181/mji.ed.204629>
- Lamidi, A. (2018). *No Title*. <https://towardsdatascience.com/breast-cancer-classification-using-support-vector-machine-svm-a510907d4878>
- Mckinney, W. (2018). *Python for Data Analysis* (2nd editio). O Reilly Media, Inc.
- Min-Wei, H., Chih-Wen, C., Wei-Chao, L., Shih-Wen, K., & Chih-Fong, T. (2018). SVM and SVM Ensembles in Breast Cancer Prediction. *Plos One*, 12. <https://doi.org/10.1371/journal.pone.0161501>
- Ortiz-Ospina, M. R. H. R. E. (2020). *Coronavirus Disease (COVID-19) – Statistics and Research*. <https://ourworldindata.org/coronavirus>

- Pambuccian, S. E. (2020). The COVID-19 pandemic: Implications for the cytology laboratory. *Journal of the American Society of Cytopathology*. <https://doi.org/https://doi.org/10.1016/j.jasc.2020.03.001>
- Permai, S. D., & Tanty, H. (2018). Linear regression model using bayesian approach for energy performance of residential building. *Procedia Computer Science*, *135*, 671–677. <https://doi.org/10.1016/j.procs.2018.08.219>
- Prem, K., Liu, Y., Russell, T., Kucharski, A. J., Eggo, R. M., Davies, N., Group, C. for the M. M. of I. D. C.-19 W., Jit, M., & Klepac, P. (2020). The effect of control strategies that reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China. *MedRxiv*, *2667(20)*, 2020.03.09.20033050. <https://doi.org/10.1101/2020.03.09.20033050>
- Qian, H. (2018). Big data Bayesian linear regression and variable selection by normal-inverse-gamma summation. *Bayesian Analysis*, *13(4)*, 1007–1031. <https://doi.org/10.1214/17-BA1083>
- Web1. <https://github.com/CSSEGISandData/COVID-19>



File1

ORIGINALITY REPORT

2%

SIMILARITY INDEX

1%

INTERNET SOURCES

0%

PUBLICATIONS

1%

STUDENT PAPERS

PRIMARY SOURCES

1

Submitted to Universitas Brawijaya

Student Paper

1%

2

pt.scribd.com

Internet Source

1%

3

www.itb.ac.id

Internet Source

<1%

4

docobook.com

Internet Source

<1%

Exclude quotes On

Exclude matches Off

Exclude bibliography On