RESEARCH ARTICLE

**Prediksi kasus Tuberculosis Menggunakan Machine Learning Python dan Algoritma Arima**

Muhammad Resha1, Anshar Suyuti2, Rhiza S Sadjad3, Muhammad Niswar4

1Department of Digital Business, IBK Nitro, Makasxsar, Indonesia  
2,3,4 Department of Electrical Engineering, Faculty of Engineering, Hasanuddin University, Makassar, Indonesia

Corresponding author. [E-mail:](mailto:amin.edalatmanesh@gmail.com) resha@nitromks.ac.id

Abstract

**BACKGROUND** :

**METHODS** :

**RESULTS** :

**CONCLUSION** :

**KEYWORDS** :

Introduction

Penyakit Tuberkulosis (TB) merupakan penyakit menular dan banyak kasus bersifat mematikan. Penyakit ini disebabkan oleh Mycobacterium Tuberculosis . Penyakit Tuberkulosis menyerang paru-paru dan dapat berdampak pada bagian tubuh lainnya. Tuberkulosis menyebar melalui udara ketika seseorang dengan infeksi TB aktif batuk, bersin, atau menyebarkan butiran ludah melalui udara. [1]

Penyakit TB masih menjadi masalah penting bagi kesehatan dunia. Dimana diperkirakan TB telah menyerang 1/3 dari penduduk dunia.[2] Data WHO pada tahun 2019 menunjukkan bahwa penyakit Tb telah menyerang sebanyak 10,4 juta jiwa dimana 1,8 juta diantaranya meninggal karena penyakitTb. WHO menyatakan bahwa 60% dari total kasus Tb di dunia berasal dari 6 negara, yaitu China, India, Indonesia, Nigeria, Pakistan, dan Afrika Selatan. [3]

Penyakit ini dapat diprediksi dengan berbagai macam model peramalan kejadian TB. Model peramalan ini dapat menjadi salah satu alat penting untuk memprediksi kejadian penyakit menular dan membantuk untuk mengambil tindakan pencegahan yang lebih cepat[4].

Memprediksi kasus TB dapat membantu mencegah penyebaran penyakit dengan deteksi dini oleh Dinas Kesehatan dan instansi terkait. Model peramalan ini dapat membantu menentukan tindakan selanjutnya dalam memberantas penyakit TB.

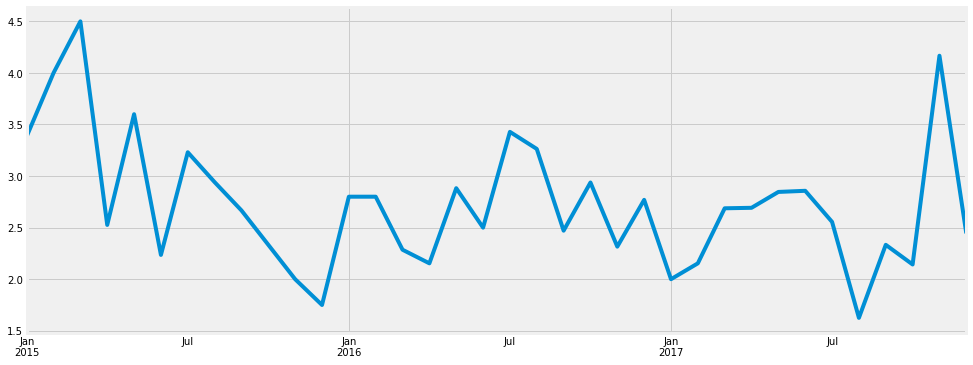
Terdapat banyak metode statistik yang dapat digunakan untuk meramalkan penyakit menular, seperti regression dan exponential smoothing, autoregressive integrated moving average (ARIMA) dan support vector machine (SVM) [5]

Model ini memperkirakan bahwa data saat ini ada hubungan linier dengan data masa lalu dan kesalahan masa lalu dari serangkaian waktu. Namun, masalah dunia nyata menunjukkan ketidakpastian dan seringkali kompleks dan nonlinier, oleh karena itu pendekatan linier mungkin tidak cocok untuk kasus di mana variabel independen dan dependen berada dalam hubungan nonlinear

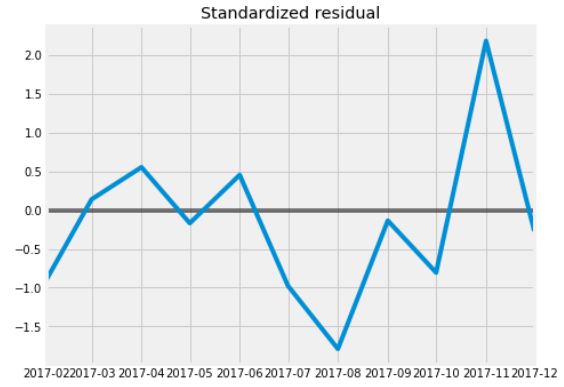
Methods

Metode prediksi time series menggunakan *Auto Regressive Integrated Moving Average* (ARIMA), yaitu metode yang menggunakan pendekatan kuantitatif masa lalu dikumpulkan dan dijadikan acuan untuk peramalan masa depan.[6]

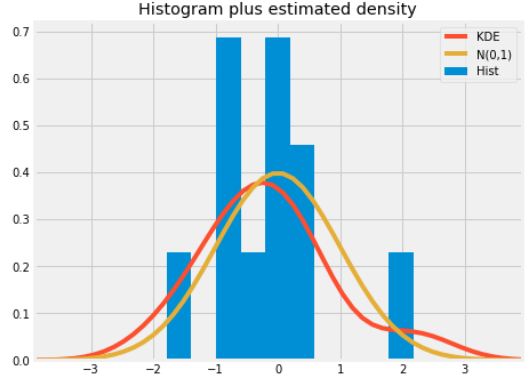
Results



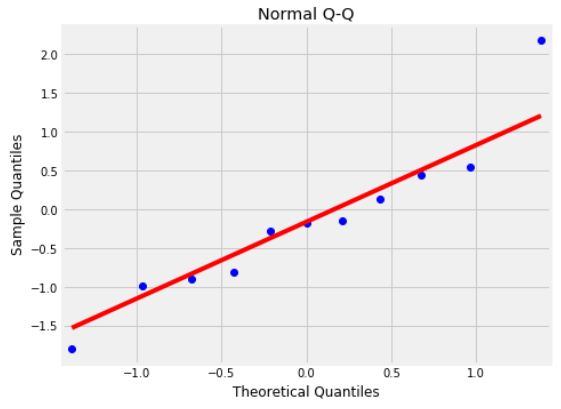
Gambar 1 : Plot TB (+) berdasarkan bulan dan tahun



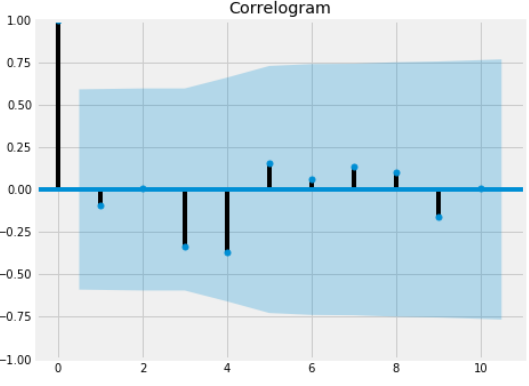
Gambar 2 : Standardized Residual



Gambar 3 : Histogram Plus Estimated Density

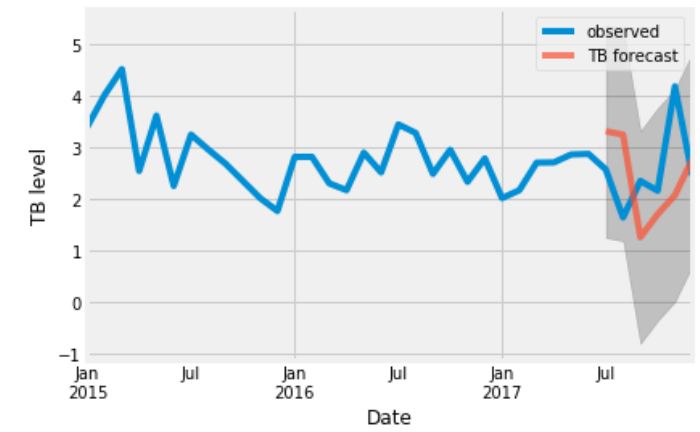


Gambar 4 : Normal Q-Q



Gambar 5 : Correlogram

Discussion



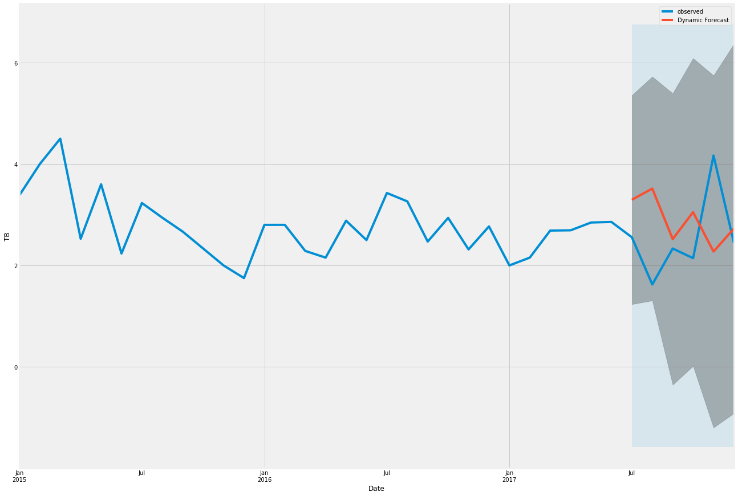
Gambar 6 : Static Forecast

Hasil prediksi dan data sebenarnya. Data di training dari tahun 2015-01-01 dan training/observasi dari tahun 2017-07

Tabel 1 : Hasil perbandingan jumlah prediksi dan jumlah sebenarnya untuk TB (+)

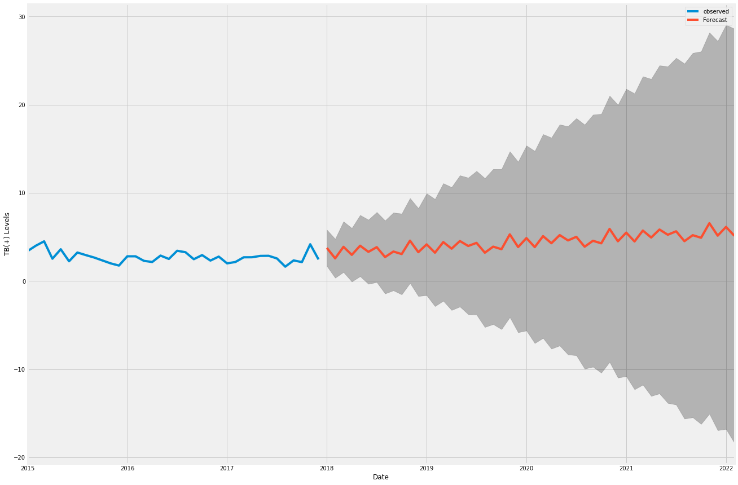
|  |  |  |
| --- | --- | --- |
| Tanggal | Jumlah prediksi | Jumlah sebenarnya |
| 01/07/2017 | 3.30 | 2.56 |
| 01/08/2017 | 3.52 | 1.63 |
| 01/09/2017 | 2.52 | 2.33 |
| 01/10/2017 | 3.05 | 2.14 |
| 01/11/2017 | 2.27 | 4.17 |
| 01/12/2017 | 2.73 | 2.44 |

The Mean Squared Error of our forecasts is 1.51



Gambar 7 : Dynamic forecast

The Mean Squared Error of our forecasts is 1.44



Gambar 8 : Forecast TB(+) untuk beberapa tahun ke depan

Tabel 2 : Hasil prediksi penyakit TB (+) hingga tahun 2022

|  |  |  |  |
| --- | --- | --- | --- |
| Tanggal | Prediksi (rata2) | Tanggal | Prediksi (rata2) |
| 01/01/2018 | 3.748706 | 01/02/2020 | 3.844061 |
| 01/02/2018 | 2.565193 | 01/03/2020 | 5.086811 |
| 01/03/2018 | 3.876617 | 01/04/2020 | 4.285079 |
| 01/04/2018 | 2.962747 | 01/05/2020 | 5.206457 |
| 01/05/2018 | 4.002225 | 01/06/2020 | 4.602238 |
| 01/06/2018 | 3.305068 | 01/07/2020 | 5.008648 |
| 01/07/2018 | 3.838346 | 01/08/2020 | 3.888259 |
| 01/08/2018 | 2.711982 | 01/09/2020 | 4.560552 |
| 01/09/2018 | 3.340983 | 01/10/2020 | 4.271238 |
| 01/10/2018 | 3.055191 | 01/11/2020 | 5.910801 |
| 01/11/2018 | 4.575247 | 01/12/2020 | 4.504095 |
| 01/12/2018 | 3.266758 | 01/01/2021 | 5.486335 |
| 01/01/2019 | 4.152521 | 01/02/2021 | 4.485857 |
| 01/02/2019 | 3.210946 | 01/03/2021 | 5.718347 |
| 01/03/2019 | 4.412886 | 01/04/2021 | 4.93096 |
| 01/04/2019 | 3.666063 | 01/05/2021 | 5.837791 |
| 01/05/2019 | 4.532347 | 01/06/2021 | 5.245097 |
| 01/06/2019 | 3.971859 | 01/07/2021 | 5.636679 |
| 01/07/2019 | 4.322973 | 01/08/2021 | 4.51806 |
| 01/08/2019 | 3.210412 | 01/09/2021 | 5.194005 |
| 01/09/2019 | 3.894713 | 01/10/2021 | 4.905091 |
| 01/10/2019 | 3.607819 | 01/11/2021 | 6.556686 |
| 01/11/2019 | 5.290032 | 01/12/2021 | 5.140093 |
| 01/12/2019 | 3.848283 | 01/01/2022 | 6.132128 |
| 01/01/2020 | 4.865349 | 01/02/2022 | 5.125762 |

Conclusion

Acknowledgments

References

[1] J. Barat, M. Rahmaniati, D. Biostatistik, and F. K. Masyarakat, “Pola Penyebaran Kasus Tuberkulosis Dengan Pendekatan Spasial-Statistik (Studi kasus di Kota Depok, Jawa Barat),” pp. 1–15, 2012.

[2] A. Zumla, A. George, V. Sharma, R. H. N. Herbert, A. Oxley, and M. Oliver, “The WHO 2014 Global tuberculosis report-further to go,” *The Lancet Global Health*, vol. 3, no. 1. Elsevier Ltd, pp. e10–e12, 01-Jan-2015.

[3] H. Getahun *et al.*, “Management of latent Mycobacterium tuberculosis infection: WHO guidelines for low tuberculosis burden countries,” *European Respiratory Journal*, vol. 46, no. 6. European Respiratory Society, pp. 1563–1576, 01-Dec-2015.

[4] K. W. Wang, C. Deng, J. P. Li, Y. Y. Zhang, X. Y. Li, and M. C. Wu, “Hybrid methodology for tuberculosis incidence time-series forecasting based on ARIMA and a NAR neural network,” *Epidemiol. Infect.*, vol. 145, no. 6, pp. 1118–1129, Apr. 2017.

[5] X. Zhang, T. Zhang, A. A. Young, and X. Li, “Applications and comparisons of four time series models in epidemiological surveillance data,” *PLoS One*, vol. 9, no. 2, Feb. 2014.

[6] D. Wiyanti and R. Pulungan, “Peramalan Deret Waktu Menggunakan Model Fungsi Basis Radial (RBF) dan Auto Regressive Integrated Moving Average (ARIMA) File Sharing System Using Samba For Mathematics Laborator View project thinning of various local character View project,” 2012.