Comparative Study of Informer, Prophet, and SARIMA Time Series Forecasting Models for Predicting Pneumonia-Related Hospitalizations and Emergency Room Visits in Elderly Patients Using OMOP-CDM

Seonghwan Shin, PharmD[1], Junhyuk Chang, PharmD[1], Min-Gyu Kim, MD[2], Byungjin Choi, MD[2], Rae Woong Park, MD, Ph.D.[1,2]
[1]Department of Biomedical Sciences, Ajou University Graduate School of Medicine, Suwon, South Korea
[2]Department of Biomedical Informatics, Ajou University School of Medicine, Suwon, South Korea

Background

- Pneumonia in elderly patients often presents fewer symptoms, making timely treatment difficult, which can lead to increased morbidity and mortality.
- As a result, sudden hospitalization and emergency room (ER) visits occur, placing a burden on healthcare resource management.
- Therefore, accurately predicting pneumonia-related hospitalizations is crucial for both patient care and efficient resource allocation.
- To address this need, this study aims to predict the daily number of pneumonia-related hospitalizations in the elderly using Prophet, SARIMA, and Informer time series forecasting models.

Methods

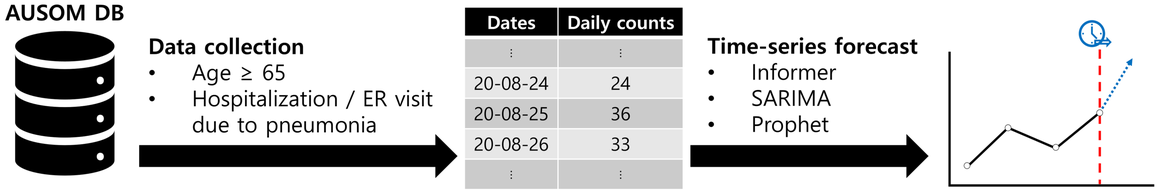

Figure 1. Framework and workflow of this study

1. Data collection

- **Database**
  - Ajou University School of Medicine (AUSOM) database (OMOP-CDM format)
- **Inclusion criteria for study population**
  - Patient records (2018-2023)
  - Age ≥ 65
  - Hospitalized or visited ER
  - Diagnosed as pneumonia within 24 hours of hospitalization or ER visit

2. Preprocessing

- Aggregated the daily counts of hospitalization and ER visits for the study population
- Missing dates are filled with 0
- Split: 80% for training / 20% for testing

3. Model development

- **Three models**
  - Prophet
  - SARIMA
  - Informer
- **Test period: 2 weeks (14 days)**
  - Compared to the actual observed counts during the test period

4. Evaluation Metrics

- Metrics used
  - Mean absolute error (MAE)
  - Root mean square error (RMSE)
- Lower metric values indicate better model performance
- Compared each model's accuracy using metrics above

Conclusion

- Informer outperformed other models.
- We confirmed the potential of advanced time series forecasting models in predicting pneumonia-related hospitalizations and ER visits in elderly patients

Contact: contact@ohdsi.org

Results

- A total of 31,338 patients, and 12,037 hospitalizations and ER visits were included.
- Informer demonstrated the lowest RMSE (1.089) and MAE (0.778), indicating superior performance.
- SARIMA followed with an RMSE of 2.595 and an MAE of 2.227.
- Prophet exhibited the highest error values, with an RMSE of 4.776 and an MAE of 4.489, reflecting the least favorable performance (Table 1, Figure 2).

Table 1. Performance metrics of the models

Models	MAE	RMSE
Informer	**0.778**	**1.089**
SARIMA	2.227	2.595
Prophet	4.489	4.776
***Note: Bold values indicate the best performance for each metric.**		

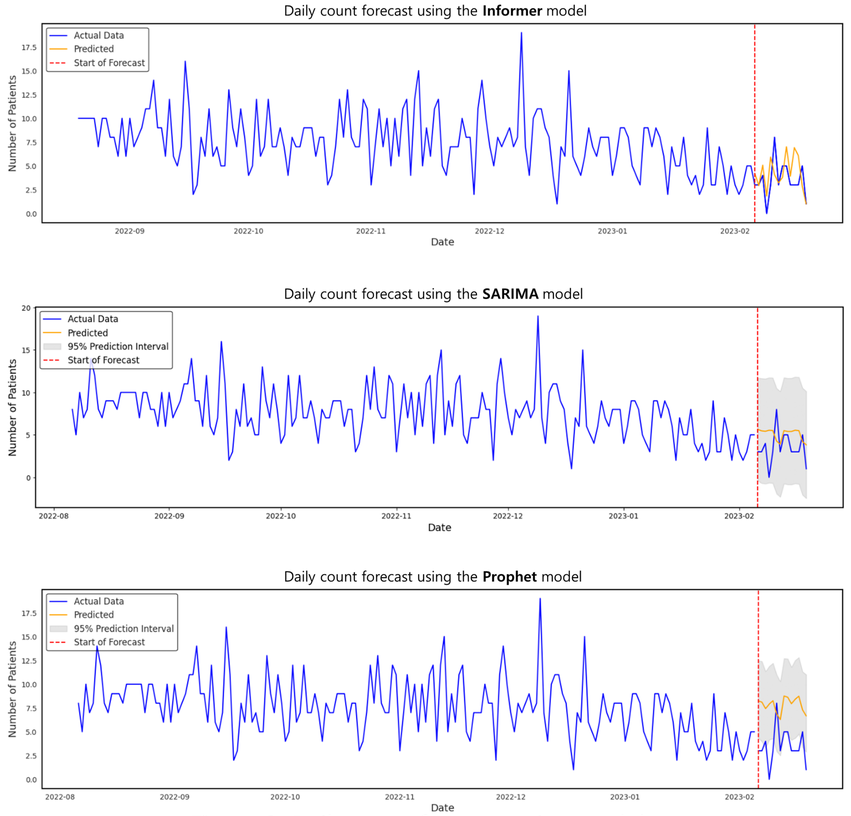

Figure 2. Daily count forecast using models

Acknowledgements

This research was funded a grant from the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number: HR16C0001) and this research was supported by a Government-wide R&D Fund project for infectious disease research (GFID), Republic of Korea (grant number: HG22C0024, KH124685).