



Beta-hCG test demand forecasting using stacking ensemble-learning and machine learning approaches

RG da Silva, V Tafoya-Martinez, FD Silva, MA Cardoso, EB Severo, CQ Cardoso, MHDM Ribeiro, VC Mariani, and LS Coelho

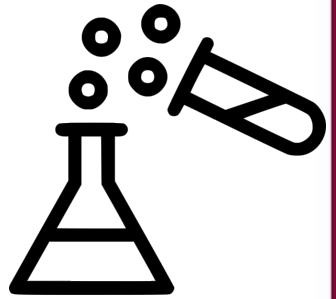
International Conference on Production Research Americas – ICPR 2022

AGENDA

- **Introduction**
- **Objective**
- **Dataset description**
- **Methodology**
- **Results**
- **Conclusion**
- **References**

INTRODUCTION

- Demand forecasting is **essential for decision-making**, since these forecasts are necessary inputs for **strategic management decisions**.
- **Hilab** is a **health-tech company** that provides laboratory tests through **Point-of-Care testing** (POCT) platform and POCT equipment using **internet of things technology** [1].
- The pregnancy test (**Beta-hCG test**) is one of the leading laboratory tests that Hilab provides.
- **Forecasting** the Beta-hCG test demand **as accurately as possible** is crucial for many areas of the company.



INTRODUCTION

- Due to these factors, forecasting sales is a **challenging task**.
- A hybrid approach that combines **Stacking ensemble-learning (STACK)** with diverse **machine learning (ML) models** can handle the time series for multi-step ahead forecasting.



OBJECTIVE

- The objective of this study is to **propose a hybrid framework to forecast the Beta-hCG test demand**, using a multi-step ahead strategy (one, fifteen, thirty, forty-five, and sixty days ahead).
- The proposed model is composed by **STACK** and **diverse ML approaches**.

DATASET DESCRIPTION

- The data consists of the **historical daily Beta-hCG test demand** from **March 11th, 2018**, to **March 31st, 2022**.
- Since the dataset is from a **private company**, it was rescaled using min-max normalization to **anonymize the actual data**.
- **The moving average of 7 days of the demand** was calculated and used as the **objective of the forecasting process**.

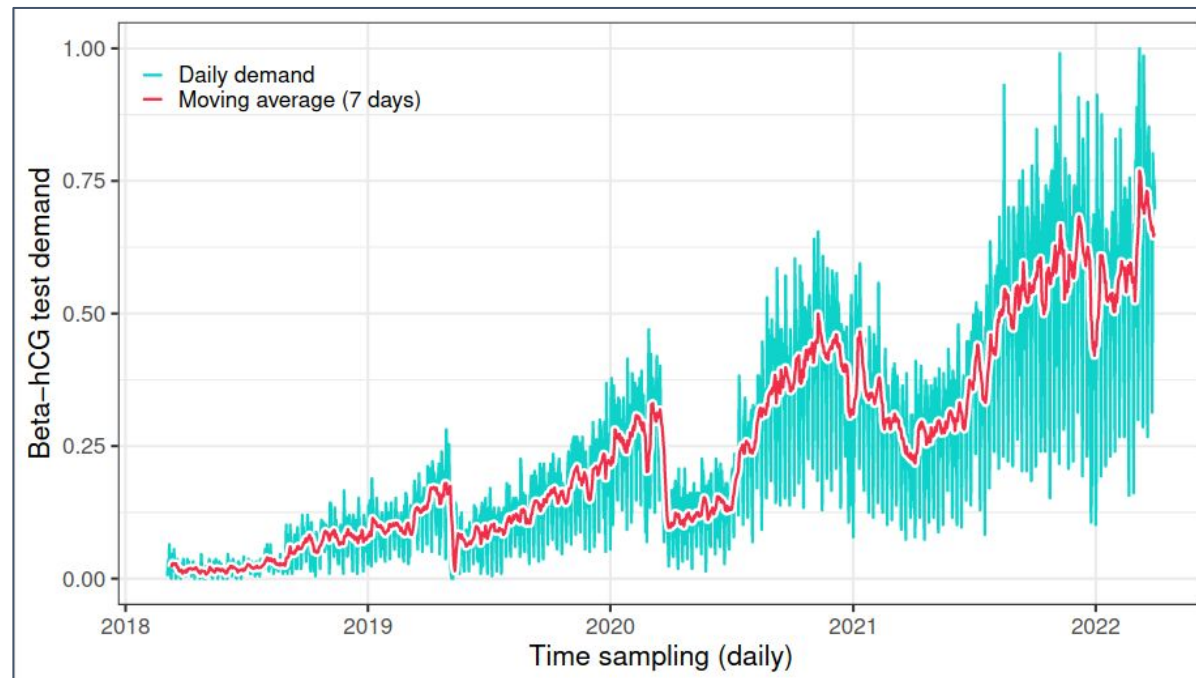


Fig. 1. Normalized time series of the Beta-hCG test demand

METHODOLOGY

- The dataset is divided into **train** and **test sets** in the **proportion of 70 and 30%**, respectively.
- **Four different algorithms** in the **STACK's layer-0**:
 - **GLM** (Generalized Linear Model).
 - **GPR** (Gaussian Process Regression).
 - **PLS** (Partial Least Squares).
 - **RIDGE** (Ridge Regression).
- **SVR** (Support Vector Regression with linear kernel) in the **STACK's layer-1**.
- Performance measures:
 - **MAE** (Mean absolute error).
 - **MAPE** (Mean absolute percentage error).
 - **RMSE** (Root mean squared error).

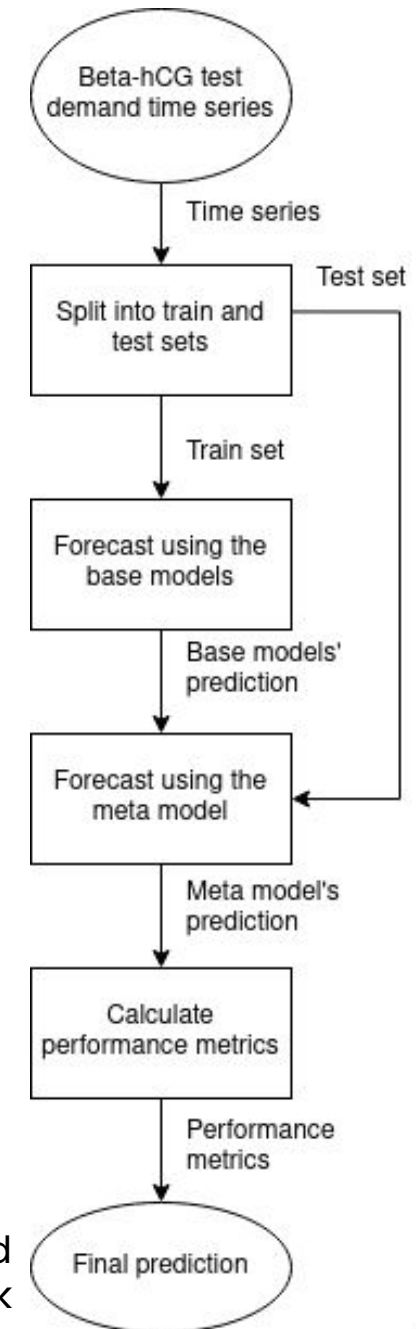


Fig. 2. Proposed forecasting framework

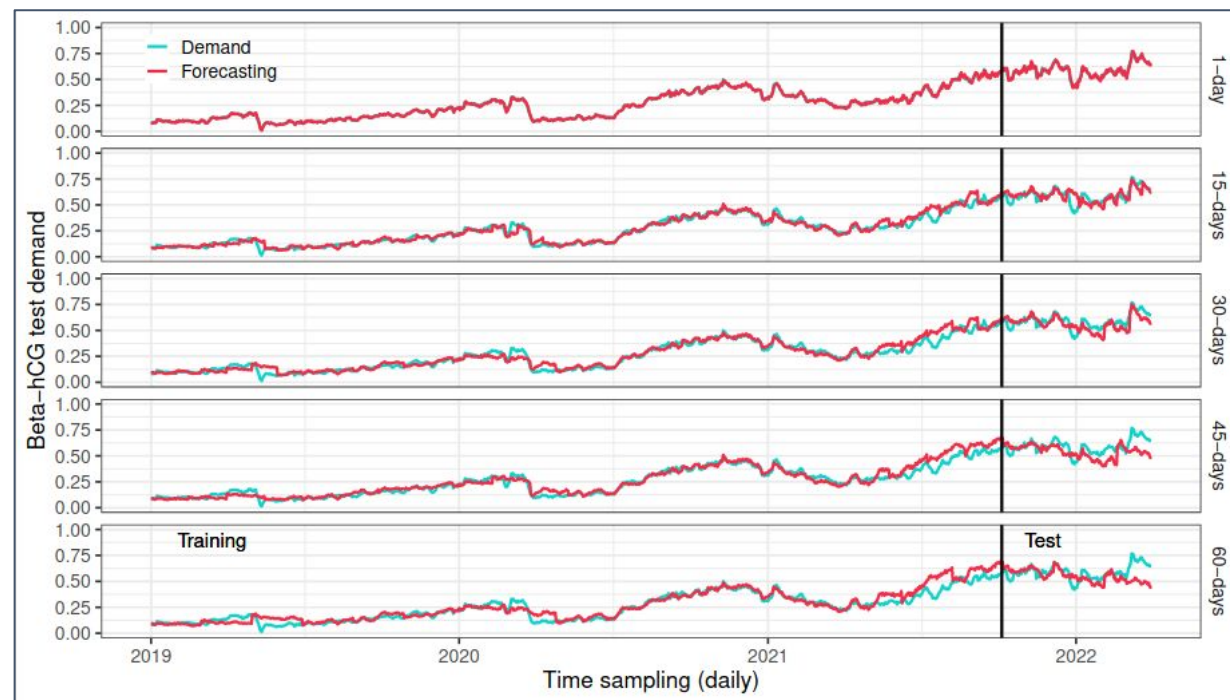
RESULTS

Table 1. Performance metrics of the compared and proposed models

Forecasting horizon	Performance metrics	GPR	GLM	PLS	RIDGE	STACK.
1 day	MAE	0.0121	0.0118	0.0330	0.0120	0.0117
	MAPE	2.53%	2.50%	6.84%	2.51%	2.47%
	RMSE	0.0162	0.0151	0.0439	0.0160	0.0150
15 days	MAE	0.0421	0.0403	0.0516	0.0422	0.0405
	MAPE	8.41%	8.08%	10.50%	8.39%	8.14%
	RMSE	0.0577	0.0559	0.0670	0.0580	0.0562
30 days	MAE	0.0532	0.0513	0.0679	0.0530	0.0501
	MAPE	10.91%	10.59%	13.76%	10.79%	10.42%
	RMSE	0.0646	0.0630	0.0813	0.0646	0.0619
45 days	MAE	0.0718	0.0707	0.0849	0.0728	0.0674
	MAPE	14.27%	14.06%	16.94%	14.26%	13.71%
	RMSE	0.0862	0.0850	0.0998	0.0867	0.0821
60 days	MAE	0.0708	0.0716	0.0816	0.0713	0.0730
	MAPE	14.25%	14.42%	16.32%	14.10%	14.52%
	RMSE	0.0916	0.0928	0.1030	0.0922	0.0942

Note: Bold values represent the best performance measure.

Fig. 3. Observed versus predicted daily Beta-hCG test demand



CONCLUSION

- This study proposed a hybrid framework by using **STACK** and **ML models** to forecast daily Beta-hCG test demanda multi-step ahead.
- The **STACK** was coupled with **GLM**, **GPR**, **PLS**, and **RIDGE** as base models and **SVR** as meta model.
- Indeed, the **STACK-based approaches** had a **better** performance than compared models in most forecasting horizons.
- **For future works**
 - Adopt different models in the layers of the STACK approach.
 - Use a signal decomposition approach to preprocess the time series.
 - Employ optimization approaches to tune hyperparameters of forecasting models.

REFERENCES

[1] Gasparin, A.T., Araujo, C.I.F., Schmitt, P., Cardoso, M.R., Perussolo, M.C., de Jesus, T.C.S., Santiago, E.B., Silva, I.L.R., de Sousa, R.G., Teng, F.Z., Severo, E.B., Ribeiro, V.H.A., Cardoso, M.A., Silva, F.D., de Araujo Perazzoli, C.R., de Holanda Farias, J.S., de Almeida, B.M.M., J´unior, S.R.R., Figueredo, M.V.M.: Hilab system, a new point-of-care hematology analyzer supported by the Internet of Things and Artificial Intelligence. Scientific Reports 12(1) (2022)

ACKNOWLEDGMENTS





Thank you!

Any questions?





Beta-hCG test demand forecasting using stacking ensemble-learning and machine learning approaches

RG da Silva, V Tafoya-Martinez, FD Silva, MA Cardoso, EB Severo, CQ Cardoso, MHDM Ribeiro, VC Mariani, and LS Coelho

International Conference on Production Research Americas – ICPR 2022