Business Analytics Online

Healthcare and Medical Analytics

Group assignment:

ICU Mortality Prediction

Group 5:

Arentsen Knud Erik Guldbaek	01766801
Kyburz Jean-Luc	01609836
Wouters Jasper	01805573
Yeung Florence	00972048
Zhang Shang	01816154

1. Introduction

Intensive Care Units (ICUs) are commonly considered expensive due to their dependence on highly trained staff and state-of-the-art technology. Recent studies show that daily, direct costs of an ICU vary between €1,168 to €2,025 (Tan SS et al, 2012; Lefrant J-Y et al, 2015), with UK costs reporting an average of £1,738 in 2016 (Marti J. et al, 2016). Though expensive, a number of empirical studies have shown that higher ICU capacity can increase the survival rates of patients having severe illness, such as cancer (Go, R.S. et al, 2016; Huo, J. et al, 2016; Birkmeyer, J.D. et al, 2007; Hillner, B.E. et al, 2000).

Given the importance and costs of scarce ICU capacity, it is of utmost importance to be able to utilize the capacity in the best possible way in order to maximise patient survival rates. In this context, accurate and timely mortality prediction methods are instrumental tools for prioritizing patients and assigning critical treatments in ICUs. Moreover, retrospective analysis of the mortality predictions can be used as foundations for ICU performance reviews and identification of areas of improvement.

Over the years, many clinical prediction tools have been developed and implemented in ICUs globally, such as APACHE II, SATS, SOFA e.t.c. Clinical prediction tools generally seem to perform well (Capuzzo et al. 2000 and Huang et al, 2017), but they are generally unusable within the initial 24 to 48 hours of patient admission to the ICU as some of the parameters are unclear at admission (Awad et al, 2020). Moreover, the clinical prediction tools to some extent rely on subjective assessment of scores, leaving room for interobserver variation in the application of scores for identical patients (Polderman et al, 2001).

As a reaction to the deficiencies of clinical predictors, recent years have seen an upspring in machine learning/data mining methods with the general ambition of utilizing objectively observable and early available patient data to generate early mortality predictions with predictive ability at least comparable to the clinical predictors. One such project is described in Van Poucke, Kovacevic and Vukicevic's paper (2018), where the authors developed a model to predict mortality up until 30 days after ICU dismissal for patients with renal failure.

The goal of this paper is to test and compare different models for predicting mortality in ICUs, using readily available vital sign measurements obtained by ICU staff as part of their ongoing observational routines. We use the MIMIC III (Medical Information Mart for Intensive Care III) dataset to test, compare and evaluate the models' ability to predict patient mortality based on observational data obtained 6, 12, and 24 hours after patient admission to the ICU.

2. Selection of predictive variables

MIMIC-III is a large and open data source that records health-related data associated with around forty thousand patients who stayed in critical care units of the Beth Israel Deaconess Medical Center between 2001 and 2012 (Johnson AEW et al, 2016). This dataset contains rich information such as demographics, vital sign measurements, laboratory test results, procedures, medications, caregiver notes, imaging reports, and mortality. In the following section, we discuss the methodology applied, and confounders controlled in this analysis.

The <u>dependent variable</u> in our analysis is a binary variable indicating whether a patient has died during his or her hospital stay. Within the MIMIC III dataset, the proportion of people that have died during their stay in the hospital is roughly 10%. The clear imbalance between the dependent variable outcomes could lead to poor performance on the minority class, which is the class we are most interested in. To remediate this shortcoming, we have oversampled the minority class using the Synthetic Minority Oversampling Technique (SMOTE) (Chawla et al, 2002).

We included <u>demographic control variables</u> available in the dataset, such as age, gender, ethnicity and the type of insurance coverage per patient. We included the type of insurance coverage to use it as a proxy for the socio-economic status per patient. Furthermore, the lack of insurance coverage can also be an indicator that a patients' illness is more severe when admitted to the ICU (O'Hara et al, 2012).

<u>Additional control variables</u> identified as factors that increase the mortality for ICU patients were included (Hong et al, 2020). Three of these variables provide information with regards to pre-existing health conditions, namely if the patient is a smoker, has diabetes or hypertension. In addition, a binary variable was included for patients that underwent surgery during their hospital stay and one numerical variable regarding the number of drugs prescribed to a patient during the hospital stay.

Lastly, a selection of <u>vital sign variables</u> were included in the final data-set. The choice was guided by 4 studies (Lee et al., 2016. Van Poucke et al, 2018. Johnson & Mark, 2017 and Alves et al, 2018.) which all tried to predict ICU mortality rates based on readily available patient vital sign data obtained within the first 48 hours of ICU admission. The 4 studies between them used 45 different variables, however with substantial overlaps. For our study, we chose to use 21 variables, which all appeared in at least 3 of the 4 studies. All variables are numeric and measured as mean values of measurements made within a given hour after ICU admission.

As we encountered missing values for some of the vital sign variables, we used the same approach as Purushotham et al (2017) to handle these values. If a feature was missing in all time windows, we used mean imputation. Otherwise, we used forward and backward imputation. Outliers above the 99% percentile were replaced using the same logic. After selecting the predictive variables, all observations in the dataset were grouped into hourly intervals. Subsequently, the data was split into 3 subsets, matching the measurements made after 6, 12 and 24 hours. The number of observations for the three subsets were 15,734, 16,511 and 15,073 respectively.

Appendix I contains the summary statistics for the entire dataset, and Appendices II, III and IV for the numeric variables in each of the 3 subsets.

3. Predictive models

As a starting point we calculated a naïve benchmark by applying a stratified strategy, using the prior probability of picking a patient with a high mortality risk, per subset.

Next, three logistic regression models were fitted (one for each of the subsets). The results of the regressions can be found in Appendices V, VI and VII.

Finally, a number of machine learning methods were applied to the dataset. Given the number of features and a binary output variable, the logistic regression and decision trees were used to generate predictive models for mortality rates in ICUs. In this report, three ensemble learning methods of decision trees were considered: Random Forest, Adaptive Boosting (Adaboost) and Gradient Boosting. Our Random Forest model utilises a bagging methodology by creating 100 decision trees and classifying each record based on the majority vote. Our Adaboost and Gradient Boosting also generate 100 decision trees, however, the former re-trains new trees based on using the same dataset but adding weights on wrong predictions and the latter creates decision trees based on the prior tree's residuals.

4. Model performance / Discussion

To compare the results of the four models, multiple metrics commonly used for binary probabilistic forecasts were calculated. Table 1 below summarizes the six metrics used for our evaluation. See Appendix VIII for the ROC-Curves and Precision-Recall Curves of the models across the three different time frames and Appendix IX for the respective confusion matrices.

Our first observation is that at least one of the predictive models are consistently able to outperform the naïve benchmark on all the performance metrics with a substantial margin.

The second observation is that the performance metrics do not seem to change significantly with the passing of time. So, the models do not become better at predicting patient mortality after 24 hours compared to the first 6 hours of ICU admission. A possible interpretation of this finding may be that the vital sign variables are already at “critical” levels at admission and the “crisis-period” in most cases lasts longer than 24 hours. Additionally, control variables like insurance type and ethnicity are highly significant and do not change in the different time windows.

Our third observation is that the machine learning methods consistently outperform the Logistic Regression on almost all performance metrics apart from the Recall score, where the Logistic Regression scores significantly higher.

The Confusion Matrixes in Appendix IX illustrate what happens here: The Logistic Regression models are consistently better at identifying True Positives, which is arguably the most important metric given that we wish to maximise the number of patients with high risk of mortality identified. However, a large number of patients may have been falsely identified as having high mortality risk, thus raising the probability that the scarce ICU capacity is spread too thinly across too many patients who do not require the treatment. To resolve this, the Average Precision scores aims to optimize the Recall and Precision for all binary probability thresholds. Here we see that the Gradient Boost model has the best performance in the first two time periods, whereas the Logistic Regression performs best after 24 hours. The ROC AUC score is a similar metric where it maximises Recall and False Positive Rates (proportion of patients wrongly labeled as going to die across all patients who live) across all thresholds. Across all three time frames, the Random Forest Model performs the best on this metric.

Metrics	Naive classifier	Logistic regression	Random forest	Adaboost	Gradient boost
T+6					
Precision	0.106	0.220	0.501	0.370	0.422
Recall	0.103	0.669	0.267	0.424	0.347
F1 Score	0.104	0.331	0.349	0.395	0.381
Accuracy	0.814	0.715	0.895	0.863	0.881
ROC AUC Score	-	0.764	0.821	0.779	0.799
Average Precision Score	-	0.355	0.381	0.365	0.383
T+12					
Precision	0.095	0.203	0.469	0.338	0.418
Recall	0.097	0.675	0.264	0.420	0.354
F1 Score	0.096	0.312	0.338	0.374	0.383
Accuracy	0.825	0.715	0.901	0.866	0.891
ROC AUC Score	-	0.776	0.817	0.795	0.807
Average Precision Score	-	0.341	0.358	0.350	0.37
T+24					
Precision	0.095	0.188	0.429	0.318	0.331
Recall	0.095	0.648	0.194	0.392	0.313
F1 Score	0.095	0.292	0.267	0.351	0.322
Accuracy	0.827	0.700	0.899	0.862	0.874
ROC AUC Score	-	0.750	0.796	0.764	0.777
Average Precision Score	-	0.311	0.296	0.297	0.304

Table 1: Performance metrics

5. Conclusion and limitations

Among the features included in the Logistic Regression model, it is interesting to note that both the insurance type and ethnicity are correlated with mortality in ICU, based on economic and statistical significances. Other important variables include age, drug prescriptions, O2 saturation pulseoxymetry, cholesterol, creatinine, respiratory rate and nitric oxide. Only 16 out of the selected 21 vital sign variables are statistically significant according to the regression.

In terms of prediction & model selection, it depends on the strategy deployed to operate the ICU facility. There is not one single model that perfectly outperforms others. The Logistic Regression models have better recall scores. But they incline to categorize more false positives. On the other hand, Random Forest models have relatively better precision scores to make sure the expensive medical resources are used efficiently. F1 score, which represents a harmonic mean of recall and precision suggests that Gradient Boosting model performs the best if we want a balanced ICU strategy. In addition, the Gradient Boosting model has the highest average precision score, second highest accuracy and ROC AUC score, based on early measures.

The analysis in this report proposes several predictive models of mortality rates and evaluates the performance of each model based on various performance metrics. However, there are limitations in the practice. Firstly, there are missing values in the MIMIC-III dataset. To mitigate its bias, we performed forward and backward imputation techniques. Secondly, oversampling by deploying SMOTE technique does not take into consideration that neighbouring examples can be from other classes. This can result in the overlapping of classes and introduce additional noise. Reducing operating errors and measuring intervals can help reduce the noise. Thirdly, the prescription and pre-existing conditions are complex factors for mortality prediction. For example, the dataset includes the number of drugs prescribed to patients but does not indicate the dosage. More refined research that focuses on specific diseases such as respiratory diseases or drug usage can provide more strategic insights. Finally, including lagged vital sign variables, providing a window on hours before time of measurement, may help to improve model performance. Further studies could address these limitations and build on our findings.

References

Alves,T., Laender, A.H., Veloso, A., & Ziviani, N. (2018). "Dynamic Prediction of ICU Mortality Risk Using Domain Adaptation*"*, 2018 *IEEE International Conference on Big Data,* p. 1328-1336.

Awad, Aya et al. (2020), "Predicting hospital mortality for intensive care unit patients: Time-series analysis", *Health Informatics Journal,* 26(2), p. 1043-1059.

Birkmeyer, J.D., Sun, Y., Wong, S.L, Stukel, T.A. (2007). "Hospital volume and late survival after cancer surgery", *Ann. Surg*. 245, p. 777–783.

Capuzzo, M et al. (2000). "Validation of severity scoring systems SAPS II and APACHE II in a single-center population", *Intensive Care Med*, 26, p. 1779-1785.

Chawla, N. V., Bowyer, K. W., Hall, L. O., and Kegelmeyer, W. P. (2002). "Smote: Synthetic minority over-sampling technique", *Journal of Artificial Intelligence Research*, 16, p. 321-357.

Go, R.S., Bartley, A.C., Crowson, C.S., Shah, N.D., Habermann, E.B., Holton, S.J., Holmes, D.R., III. (2016). "Association between treatment facility volume and mortality of patients with multiple myeloma", *J. Clin. Oncol.*, 35, p. 598–604.

Hillner, B.E., Smith, T.J., Desch, C.E. (2000). "Hospital and physician volume or specialization and outcomes in cancer treatment: Importance in quality of cancer care" *J. Clin. Oncol.* 18, p. 2327–2340.

Hong, Gang Ren et al. (2020). "*Risk Factors for ICU Admission, Mechanical Ventilation and Mortality in Hospitalized Patients with COVID-19 in Hubei*, *China*" [online]. Available at (medRxiv preprint doi): https://doi.org/10.1101/2020.08.31.20184952 (Accessed: 19 December 2020).

Huang, Jianshu et al. (2017). "The value of APCAHE II in predicting mortality after paraquat poisoning in Chinese and Korean population. A systematic review and meta analysis", *Medicine*, 96(30) [online]. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5627797/# (Accessed: 6 December 2020).

Huo, J., Lairson, D.R.; Du, X.L., Chan, W., Jiang, J., Buchholz, T.A., Guadagnolo, B.A. (2016). "Hospital case volume is associated with improved survival for patients with metastatic melanoma", *Am J Clin Oncol 2016; 39;* p. 491-496.

Johnson A.E.W., Pollard T.J., Shen L., Lehman L., Feng M., Ghassemi M., Moody B., Szolovits P., Celi L.A., and Mark R.G., (2016). "MIMIC-III, a freely accessible critical care database", *Scientific Data*, [online]. Available at: http://www.nature.com/articles/sdata201635 (Accessed: 12 December 2020).

Johnson, A., Mark, R. (2017). "Real-time mortality predictions in the Intensive Care Unit", *AMIA Annual Symposium Proceedings,* p. 994-1003.

Lee, J. et al (2016). "Mortality Prediction in the ICU". Chapter 21 in: "Secondary Analysis of Electronic Health Records", *MIT Critical Data,* Springer.

Lefrant J-Y., Garrigues B., Pribil C., Bardoulat I., Courtial F., Maurel F., et al. (2015), "The daily cost of ICU patients: a micro-costing study in 23 French Intensive Care Units", *Anaesth Crit Care Pain Med*, 34, p. 151–157.

Marti J., Hall P., Hamilton P., Lamb S., McCabe C., Lall R., et al. (2016). "One-year resource utilisation, costs and quality of life in patients with acute respiratory distress syndrome (ARDS): secondary analysis of a randomised controlled trial", *Journal of Intensive Care* ,4 [online]. Available at: https://doi.org/10.1186/s40560-016-0178-8 (Accessed: 11 December 2020).

O'Hara, B., Caswell, K. (2012). "Health Status, Health Insurance, and Medical Services Utilization: 2010," *Current Population Reports,* U.S. Census Bureau, Washington, DC.

Purushotham, S., Meng, C., Che, Z., Liu, Y. (2017). "Benchmark of deep learning on models on large healthcare MIMIC datasets" [online]. Available at: https://arxiv.org/pdf/1710.08531.pdf (Accessed: 12 December 2020).

Polderman, K.H et al (2001). "Accuracy and Reliability of APACHE II scoring in two intensive cate units", *Anaesthesia*, 56, p. 47-50.

Saito, T. and Rehmsmeier, M. (2015). "The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets" PLOS ONE, [online] 10(3). Available at: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0118432 (Accessed 25 December 2020).

Tan S.S., Bakker J., Hoogendoorn M.E., Kapila A., Martin J., Pezzi A., et al. (2012). "Direct cost analysis of intensive care unit stay in four European countries: applying a standardized costing methodology", *Value Health*, 15, p. 81–86.

Poucke, S.V., Kovacevic, A. and Vukicevic, M. (2018). “Early Prediction of Patient Mortality Based on Routine Laboratory Tests and Predictive Models in Critically Ill Patients” in Ciza. T. (ed.) *Data Mining*. London: IntechOpen, p. 93-106.

Appendix I: General characteristics of the study population (summary)

Variable	N	Mean	SD	SE	95% Conf. Interval	
Male	383,413.00	0.56	0.50	0.00	0.56	0.56
Age	383,413.00	63.94	17.75	0.03	63.88	63.99
Smoke	383,413.00	0.03	0.17	0.00	0.03	0.03
Diabetes	383,413.00	0.16	0.36	0.00	0.16	0.16
Hypertension	383,413.00	0.25	0.43	0.00	0.25	0.25
Surgery	383,413.00	0.05	0.22	0.00	0.05	0.05
Drugs	383,413.00	1.24	6.71	0.01	1.21	1.26
Heart Rate	383,413.00	84.33	20.02	0.03	84.27	84.39
Non Invasive Blood Pressure mean	383,413.00	76.47	15.11	0.02	76.42	76.52
Hemoglobin	383,413.00	10.82	1.74	0.00	10.81	10.82
O2 saturation pulseoxymetry	383,413.00	97.00	3.60	0.01	96.99	97.01
WBC	383,413.00	10.91	5.25	0.01	10.89	10.93
Cholesterol	383,413.00	156.99	9.91	0.02	156.96	157.02
Creatinine	383,413.00	1.21	0.83	0.00	1.21	1.22
Magnesium	383,413.00	1.99	0.37	0.00	1.99	1.99
Temperature Fahrenheit	383,413.00	98.15	1.37	0.00	98.14	98.15
Respiratory Rate (Total)	383,413.00	18.55	3.31	0.01	18.54	18.57
Nitric Oxide	383,413.00	26.24	0.26	0.00	26.24	26.24
Arterial Base Excess	383,413.00	-0.98	2.59	0.00	-0.99	-0.98
Ionized Calcium	383,413.00	1.14	0.82	0.00	1.14	1.14
Lactic Acid	383,413.00	2.07	1.10	0.00	2.07	2.07
Phosphorous	383,413.00	3.45	0.95	0.00	3.45	3.45
Total Bilirubin	383,413.00	1.21	0.93	0.00	1.21	1.22
Sodium (whole blood)	383,413.00	137.01	2.31	0.00	137.01	137.02
Chloride (whole blood)	383,413.00	106.68	2.55	0.00	106.67	106.69
Glucose (whole blood)	383,413.00	142.33	28.98	0.05	142.24	142.42
Hematocrit (whole blood - calc)	383,413.00	30.15	2.13	0.00	30.14	30.16
Anion gap	383,413.00	13.46	3.04	0.00	13.45	13.47
Death	383,413.00	0.10	0.30	0.00	0.10	0.10

Table 2: Summary statistics - Numeric

Variable	Outcome	Count	Percent
Insurance	Medicare	212,469.00	55.42
	Private	113,969.00	29.72
	Medicaid	40,547.00	10.58
	Government	12,546.00	3.27
	Self Pay	3,882.00	1.01
Ethnicity	White	276,146.00	72.02
	Black / African American	42,826.00	11.17
	Other / Unknown	38,187.00	9.96
	Hispanic / Latino	15,976.00	4.17
	Asian	10,278.00	2.68

Table 3: Summary statistics - Categorical

Appendix II: Descriptive statistics for numeric variables (T+6)

	count	mean	std	min	25%	50%	75%	max
Diabetes	15,734.00	0.15	0.36	0.00	0.00	0.00	0.00	1.00
Hypertension	15,734.00	0.25	0.43	0.00	0.00	0.00	0.00	1.00
Surgery	15,734.00	0.04	0.21	0.00	0.00	0.00	0.00	1.00
Drug	15,734.00	0.96	5.50	0.00	0.00	0.00	0.00	140.00
Heart Rate	15,734.00	84.08	18.21	0.00	71.00	83.00	95.00	174.50
Non Invasive Blood Pressure mean	15,734.00	75.99	15.04	0.00	66.00	76.00	84.00	203.00
Hemoglobin	15,734.00	10.85	1.78	6.70	9.60	10.82	12.10	14.70
O2 saturation pulseoxymetry	15,734.00	97.19	3.15	9.00	96.00	98.00	100.00	100.00
WBC	15,734.00	10.97	5.37	0.10	7.30	10.20	13.50	33.80
Cholesterol	15,734.00	157.02	10.08	50.00	156.99	156.99	156.99	273.00
Creatinine	15,734.00	1.22	0.84	0.00	0.70	1.00	1.30	5.20
Magnesium	15,734.00	1.96	0.38	0.20	1.70	1.99	2.10	6.50
Temperature Fahrenheit	15,734.00	98.02	1.38	86.30	97.20	98.00	98.70	106.40
Respiratory Rate (Total)	15,734.00	18.50	3.24	0.00	18.55	18.55	18.55	47.00
Nitric Oxide	15,734.00	26.25	0.20	20.00	26.24	26.24	26.24	40.00
Arterial Base Excess	15,734.00	-1.07	2.65	-12.00	-0.98	-0.98	-0.98	8.00
Ionized Calcium	15,734.00	1.14	0.83	0.25	1.14	1.14	1.14	105.00
Lactic Acid	15,734.00	2.13	1.16	0.30	1.80	2.07	2.07	21.40
Phosphorous	15,734.00	3.45	0.95	0.80	2.90	3.45	3.90	6.50
Total Bilirubin	15,734.00	1.20	0.90	0.00	0.90	1.21	1.21	9.60
Sodium (whole blood)	15,734.00	137.01	2.35	0.00	137.01	137.01	137.01	177.00
Chloride (whole blood)	15,734.00	106.69	2.55	3.40	106.68	106.68	106.68	149.00
Glucose (whole blood)	15,734.00	142.98	29.45	28.00	142.33	142.33	142.33	786.00
Hematocrit (whole blood - calc)	15,734.00	30.13	2.10	17.90	30.15	30.15	30.15	42.00
Anion gap	15,734.00	13.65	3.12	6.00	11.00	13.46	15.00	22.00
Death	15,734.00	0.10	0.30	0.00	0.00	0.00	0.00	1.00

Table 4: Descriptive statistics - Numeric (T+6)

Appendix III: Descriptive statistics for numeric variables (T+12)

	count	mean	std	min	25%	50%	75%	max
Diabetes	16,511.00	0.16	0.36	0.00	0.00	0.00	0.00	1.00
Hypertension	16,511.00	0.25	0.43	0.00	0.00	0.00	1.00	1.00
Surgery	16,511.00	0.05	0.22	0.00	0.00	0.00	0.00	1.00
Drug	16,511.00	1.16	6.13	0.00	0.00	0.00	0.00	220.00
Heart Rate	16,511.00	83.73	17.80	29.00	71.00	83.00	95.00	194.78
Non Invasive Blood Pressure mean	16,511.00	75.81	14.73	0.00	66.00	76.00	83.00	204.00
Hemoglobin	16,511.00	10.84	1.74	6.70	9.60	10.82	12.10	14.70
O2 saturation pulseoxymetry	16,511.00	97.00	3.10	0.00	96.00	97.00	99.00	100.00
WBC	16,511.00	10.88	5.26	0.10	7.25	10.10	13.30	33.80
Cholesterol	16,511.00	156.99	9.91	50.00	156.99	156.99	156.99	273.00
Creatinine	16,511.00	1.21	0.83	0.00	0.70	1.00	1.30	5.20
Magnesium	16,511.00	1.99	0.37	0.70	1.80	1.99	2.20	6.60
Temperature Fahrenheit	16,511.00	98.18	1.32	88.88	97.40	98.15	98.90	105.08
Respiratory Rate (Total)	16,511.00	18.55	3.25	-1.00	18.55	18.55	18.55	45.00
Nitric Oxide	16,511.00	26.24	0.24	5.00	26.24	26.24	26.24	40.00
Arterial Base Excess	16,511.00	-1.01	2.55	-12.00	-0.98	-0.98	-0.98	8.50
Ionized Calcium	16,511.00	1.14	0.81	0.25	1.14	1.14	1.14	105.00
Lactic Acid	16,511.00	2.07	1.07	0.20	1.70	2.07	2.07	19.80
Phosphorous	16,511.00	3.46	0.94	0.80	2.90	3.45	3.90	6.50
Total Bilirubin	16,511.00	1.21	0.92	0.00	1.00	1.21	1.21	9.60
Sodium (whole blood)	16,511.00	137.02	2.36	0.00	137.01	137.01	137.01	176.00
Chloride (whole blood)	16,511.00	106.68	2.53	3.40	106.68	106.68	106.68	149.00
Glucose (whole blood)	16,511.00	142.35	28.05	0.00	142.33	142.33	142.33	786.00
Hematocrit (whole blood - calc)	16,511.00	30.16	2.10	17.90	30.15	30.15	30.15	42.00
Anion gap	16,511.00	13.44	3.04	6.00	11.00	13.00	15.00	22.00
Death	16,511.00	0.10	0.30	0.00	0.00	0.00	0.00	1.00

Table 5: Descriptive statistics - Numeric (T+12)

Appendix IV: Descriptive statistics for numeric variables (T+24)

	count	mean	std	min	25%	50%	75%	max
Diabetes	14,819.00	0.16	0.36	0.00	0.00	0.00	0.00	1.00
Hypertension	14,819.00	0.25	0.43	0.00	0.00	0.00	0.00	1.00
Surgery	14,819.00	0.06	0.24	0.00	0.00	0.00	0.00	1.00
Drug	14,819.00	1.82	8.70	0.00	0.00	0.00	0.00	148.00
Heart Rate	14,819.00	84.23	17.20	0.00	72.00	83.00	95.00	166.00
Non Invasive Blood Pressure mean	14,819.00	75.96	14.49	0.00	67.00	76.47	83.00	203.00
Hemoglobin	14,819.00	10.73	1.66	6.70	9.50	10.80	11.90	14.70
O2 saturation pulseoxymetry	14,819.00	96.76	3.15	0.00	95.00	97.00	99.00	100.00
WBC	14,819.00	10.97	5.05	0.10	7.60	10.40	13.30	33.80
Cholesterol	14,819.00	156.94	9.59	50.00	156.99	156.99	156.99	273.00
Creatinine	14,819.00	1.21	0.82	0.00	0.70	1.00	1.30	5.20
Magnesium	14,819.00	2.03	0.35	0.80	1.80	2.00	2.20	6.40
Temperature Fahrenheit	14,819.00	98.31	1.29	88.88	97.60	98.20	99.00	105.08
Respiratory Rate (Total)	14,819.00	18.51	3.52	0.00	18.55	18.55	18.55	45.00
Nitric Oxide	14,819.00	26.24	0.40	0.00	26.24	26.24	26.24	40.00
Arterial Base Excess	14,819.00	-0.84	2.54	-12.00	-0.98	-0.98	0.00	8.50
Ionized Calcium	14,819.00	1.14	0.86	0.59	1.14	1.14	1.14	105.00
Lactic Acid	14,819.00	1.99	0.98	0.20	1.60	2.07	2.07	18.80
Phosphorous	14,819.00	3.43	0.94	0.80	2.90	3.45	3.90	6.50
Total Bilirubin	14,819.00	1.23	0.96	0.00	1.00	1.21	1.21	9.60
Sodium (whole blood)	14,819.00	137.01	2.41	0.00	137.01	137.01	137.01	177.00
Chloride (whole blood)	14,819.00	106.68	2.62	3.40	106.68	106.68	106.68	149.00
Glucose (whole blood)	14,819.00	140.82	28.18	0.00	142.33	142.33	142.33	704.00
Hematocrit (whole blood - calc)	14,819.00	30.11	2.26	17.90	30.15	30.15	30.15	42.33
Anion gap	14,819.00	13.18	2.94	6.00	11.00	13.00	15.00	22.00
Death	14,819.00	0.10	0.29	0.00	0.00	0.00	0.00	1.00

Table 6: Descriptive statistics - Numeric (T+24)

Appendix V: Logistic regression (T+6)

Variable	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
Insurance - Government	0.71	0.03	22.67	0.00	0.64	0.77
Insurance - Medicaid	0.09	0.02	4.33	0.00	0.05	0.13
Insurance - Medicare	0.05	0.02	2.36	0.02	0.01	0.09
Insurance - Private	-0.10	0.02	-4.76	0.00	-0.14	-0.06
Insurance - Self Pay	-0.10	0.02	-4.93	0.00	-0.14	-0.06
Ethnicity - Asian	-0.12	0.02	-6.00	0.00	-0.16	-0.08
Ethnicity - Black/African	0.19	0.02	8.58	0.00	0.14	0.23
Ethnicity - Hispanic/Latino	0.32	0.02	14.23	0.00	0.27	0.36
Ethnicity - Other	-0.02	0.02	-0.88	0.38	-0.06	0.02
Ethnicity - White	-0.18	0.02	-7.52	0.00	-0.23	-0.13
Male	-0.06	0.02	-3.07	0.00	-0.09	-0.02
Age	0.21	0.02	9.99	0.00	0.17	0.25
Smoke	-0.01	0.02	-0.42	0.67	-0.05	0.03
Diabetes	0.06	0.02	2.54	0.01	0.01	0.11
Hypertension	-0.04	0.02	-1.97	0.05	-0.09	0.00
Surgery	-0.22	0.02	-10.83	0.00	-0.26	-0.18
Drugs	0.20	0.02	10.53	0.00	0.17	0.24
Heart Rate	-0.04	0.04	-1.02	0.31	-0.12	0.04
Non Invasive Blood Pressure mean	0.08	0.02	3.57	0.00	0.03	0.12
Hemoglobin	-2.71	0.39	-7.03	0.00	-3.47	-1.96
O2 saturation pulseoxymetry	0.32	0.02	12.90	0.00	0.27	0.36
WBC	0.09	0.02	3.82	0.00	0.04	0.13
Cholesterol	0.08	0.02	4.24	0.00	0.04	0.12
Creatinine	0.12	0.02	5.27	0.00	0.08	0.17
Magnesium	0.04	0.03	1.67	0.09	-0.01	0.09
Temperature Fahrenheit	0.00	0.02	0.21	0.83	-0.04	0.05
Respiratory Rate (Total)	0.12	0.02	5.48	0.00	0.08	0.16
Nitric Oxide	0.36	0.02	14.49	0.00	0.31	0.41
Arterial Base Excess	-0.93	1,033,830.00	0.00	1.00	-2,026,270.00	2,026,268.00
Ionized Calcium	-0.15	1,033,830.00	0.00	1.00	-2,026,269.00	2,026,269.00
Lactic Acid	-0.09	1,033,830.00	0.00	1.00	-2,026,269.00	2,026,269.00
Phosphorous	-0.23	1,033,830.00	0.00	1.00	-2,026,269.00	2,026,269.00
Total Bilirubin	-0.41	1,033,830.00	0.00	1.00	-2,026,270.00	2,026,269.00
Sodium (whole blood)	-0.11	1,033,830.00	0.00	1.00	-2,026,269.00	2,026,269.00
Chloride (whole blood)	-1.12	1,033,830.00	0.00	1.00	-2,026,270.00	2,026,268.00
Glucose (whole blood)	-0.43	1,033,830.00	0.00	1.00	-2,026,270.00	2,026,269.00
Hematocrit (whole blood - calc)	0.15	1,033,830.00	0.00	1.00	-2,026,269.00	2,026,269.00
Anion gap	-0.31	1,033,830.00	0.00	1.00	-2,026,269.00	2,026,269.00

Table 7: Logistic Regression (T+6)

Model:	Logit
No. Observations:	14,136.0
Df Model:	36.0
Df Residuals:	14,099.0
Converged:	1.0
No. Iterations:	85.0
Pseudo R-squared:	0.2
AIC:	15,722.3
BIC:	16,001.9
Log-Likelihood:	-7,824.1
LL-Null:	-9,798.3
LLR p-value:	0.0
Scale:	1.0

Appendix VI: Logistic regression (T+12)

Variable	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
Insurance - Government	0.66	0.03	22.95	0.00	0.61	0.72
Insurance - Medicaid	0.06	0.02	2.93	0.00	0.02	0.10
Insurance - Medicare	-0.02	0.02	-0.97	0.33	-0.06	0.02
Insurance - Private	-0.13	0.02	-6.06	0.00	-0.17	-0.08
Insurance - Self Pay	-0.16	0.02	-8.03	0.00	-0.20	-0.12
Ethnicity - Asian	-0.28	0.02	-11.33	0.00	-0.33	-0.23
Ethnicity - Black/African	0.19	0.02	9.23	0.00	0.15	0.23
Ethnicity - Hispanic/Latino	0.19	0.02	8.95	0.00	0.15	0.23
Ethnicity - Other	-0.11	0.02	-5.28	0.00	-0.15	-0.07
Ethnicity - White	-0.17	0.02	-7.87	0.00	-0.22	-0.13
Male	0.03	0.02	1.65	0.10	-0.01	0.07
Age	0.28	0.02	13.39	0.00	0.24	0.33
Smoke	-0.03	0.02	-1.75	0.08	-0.07	0.00
Diabetes	0.02	0.02	0.93	0.35	-0.02	0.07
Hypertension	0.02	0.02	0.73	0.46	-0.03	0.06
Surgery	-0.04	0.02	-1.87	0.06	-0.07	0.00
Drugs	0.26	0.02	13.57	0.00	0.22	0.30
Heart Rate	0.01	0.04	0.42	0.68	-0.05	0.08
Non Invasive Blood Pressure mean	-0.05	0.02	-2.24	0.03	-0.09	-0.01
Hemoglobin	-0.18	0.02	-9.08	0.00	-0.22	-0.14
O2 saturation pulseoxymetry	0.25	0.02	10.75	0.00	0.20	0.30
WBC	-0.03	0.02	-1.19	0.23	-0.07	0.02
Cholesterol	0.12	0.02	6.57	0.00	0.08	0.16
Creatinine	0.15	0.02	6.90	0.00	0.11	0.19
Magnesium	-0.07	0.02	-2.86	0.00	-0.12	-0.02
Temperature Fahrenheit	0.06	0.02	2.83	0.00	0.02	0.10
Respiratory Rate (Total)	0.08	0.02	3.82	0.00	0.04	0.13
Nitric Oxide	0.35	0.02	14.99	0.00	0.30	0.39
Arterial Base Excess	-0.84	372,718.40	0.00	1.00	-730,515.49	730,513.80
Ionized Calcium	-0.16	372,718.40	0.00	1.00	-730,514.80	730,514.49
Lactic Acid	-0.09	372,718.40	0.00	1.00	-730,514.74	730,514.56
Phosphorous	-0.19	372,718.40	0.00	1.00	-730,514.84	730,514.46
Total Bilirubin	0.22	372,718.40	0.00	1.00	-730,514.43	730,514.87
Sodium (whole blood)	0.27	372,718.40	0.00	1.00	-730,514.38	730,514.91
Chloride (whole blood)	-0.96	372,718.40	0.00	1.00	-730,515.61	730,513.69
Glucose (whole blood)	-0.37	372,718.40	0.00	1.00	-730,515.02	730,514.27
Hematocrit (whole blood - calc)	0.35	372,718.40	0.00	1.00	-730,514.30	730,514.99
Anion gap	-0.34	372,718.40	0.00	1.00	-730,514.98	730,514.31

Table 8: Logistic Regression (T+12)

Model:	Logit
No. Observations:	14,892.0
Df Model:	36.0
Df Residuals:	14,855.0
Converged:	1.0
No. Iterations:	47.0
Pseudo R-squared:	0.2
AIC:	16,729.3
BIC:	17,010.8
Log-Likelihood:	-8,327.7
LL-Null:	-10,322.0
LLR p-value:	0.0
Scale:	1.0

Appendix VII: Logistic regression (T+24)

Variable	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
Insurance - Government	0.68	0.03	21.43	0.00	0.62	0.74
Insurance - Medicaid	0.07	0.02	3.33	0.00	0.03	0.11
Insurance - Medicare	0.03	0.02	1.66	0.10	-0.01	0.07
Insurance - Private	-0.18	0.02	-8.14	0.00	-0.23	-0.14
Insurance - Self Pay	-0.12	0.02	-5.70	0.00	-0.16	-0.08
Ethnicity - Asian	-0.31	0.03	-12.03	0.00	-0.36	-0.26
Ethnicity - Black/African	0.13	0.02	7.11	0.00	0.10	0.17
Ethnicity - Hispanic/Latino	0.30	0.02	13.37	0.00	0.25	0.34
Ethnicity - Other	-0.19	0.02	-7.93	0.00	-0.23	-0.14
Ethnicity - White	-0.16	0.02	-6.85	0.00	-0.21	-0.12
Male	0.04	0.02	2.35	0.02	0.01	0.08
Age	0.27	0.02	12.36	0.00	0.23	0.31
Smoke	-0.05	0.02	-2.84	0.00	-0.09	-0.02
Diabetes	0.12	0.02	4.97	0.00	0.08	0.17
Hypertension	0.05	0.02	2.43	0.02	0.01	0.09
Surgery	-0.06	0.02	-3.04	0.00	-0.10	-0.02
Drugs	0.26	0.02	12.32	0.00	0.22	0.30
Heart Rate	-0.16	0.81	-0.20	0.84	-1.75	1.42
Non Invasive Blood Pressure mean	-0.03	0.02	-1.60	0.11	-0.07	0.01
Hemoglobin	-0.34	0.45	-0.76	0.45	-1.21	0.54
O2 saturation pulseoxymetry	0.15	0.02	6.98	0.00	0.11	0.20
WBC	0.01	0.02	0.48	0.63	-0.04	0.06
Cholesterol	0.14	0.02	7.36	0.00	0.10	0.17
Creatinine	0.13	0.02	5.61	0.00	0.09	0.18
Magnesium	-0.07	0.03	-2.92	0.00	-0.12	-0.02
Temperature Fahrenheit	0.11	0.02	5.02	0.00	0.07	0.16
Respiratory Rate (Total)	0.09	0.02	3.61	0.00	0.04	0.13
Nitric Oxide	0.37	0.03	14.54	0.00	0.32	0.42
Arterial Base Excess	-0.62	645,884.56	0.00	1.00	-1,265,911.00	1,265,910.00
Ionized Calcium	-0.49	645,884.56	0.00	1.00	-1,265,911.00	1,265,910.00
Lactic Acid	-0.28	645,884.56	0.00	1.00	-1,265,911.00	1,265,910.00
Phosphorous	-0.34	645,884.56	0.00	1.00	-1,265,911.00	1,265,910.00
Total Bilirubin	0.47	645,884.56	0.00	1.00	-1,265,910.00	1,265,911.00
Sodium (whole blood)	-0.34	645,884.56	0.00	1.00	-1,265,911.00	1,265,910.00
Chloride (whole blood)	-0.90	645,884.56	0.00	1.00	-1,265,911.00	1,265,910.00
Glucose (whole blood)	-0.24	645,884.56	0.00	1.00	-1,265,911.00	1,265,910.00
Hematocrit (whole blood - calc)	0.36	645,884.56	0.00	1.00	-1,265,910.00	1,265,911.00
Anion gap	-0.15	645,884.56	0.00	1.00	-1,265,911.00	1,265,910.00

Table 9: Logistic Regression (T+24)

Model:	Logit
No. Observations:	13,402.0
Df Model:	36.0
Df Residuals:	13,365.0
Converged:	1.0
No. Iterations:	70.0
Pseudo R-squared:	0.2
AIC:	14,983.3
BIC:	15,260.9
Log-Likelihood:	-7,454.6
LL-Null:	-9,289.6
LLR p-value:	0.0
Scale:	1.0

Appendix VIII: Model Performance

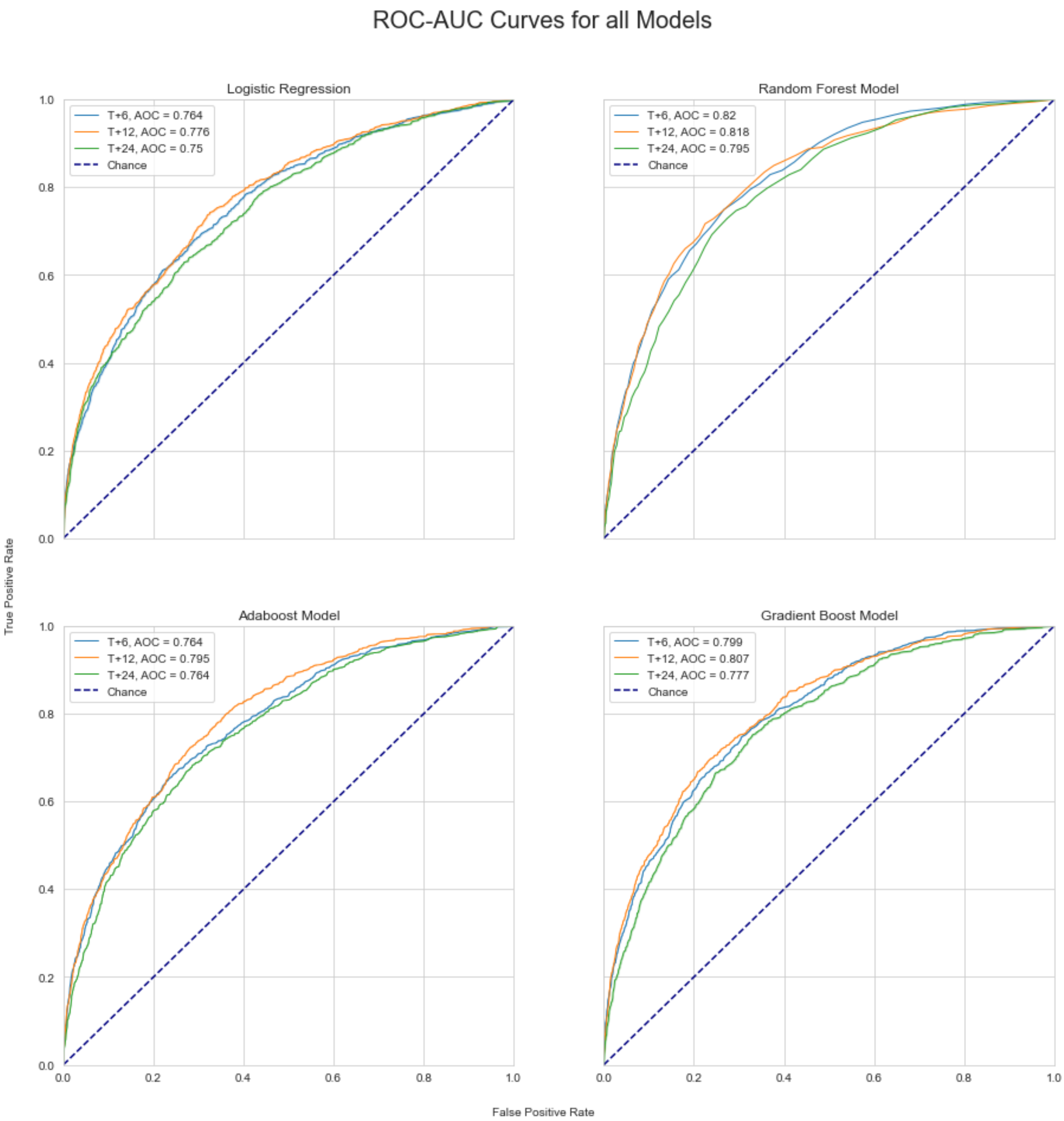

Figure 1: ROC-AUC curve

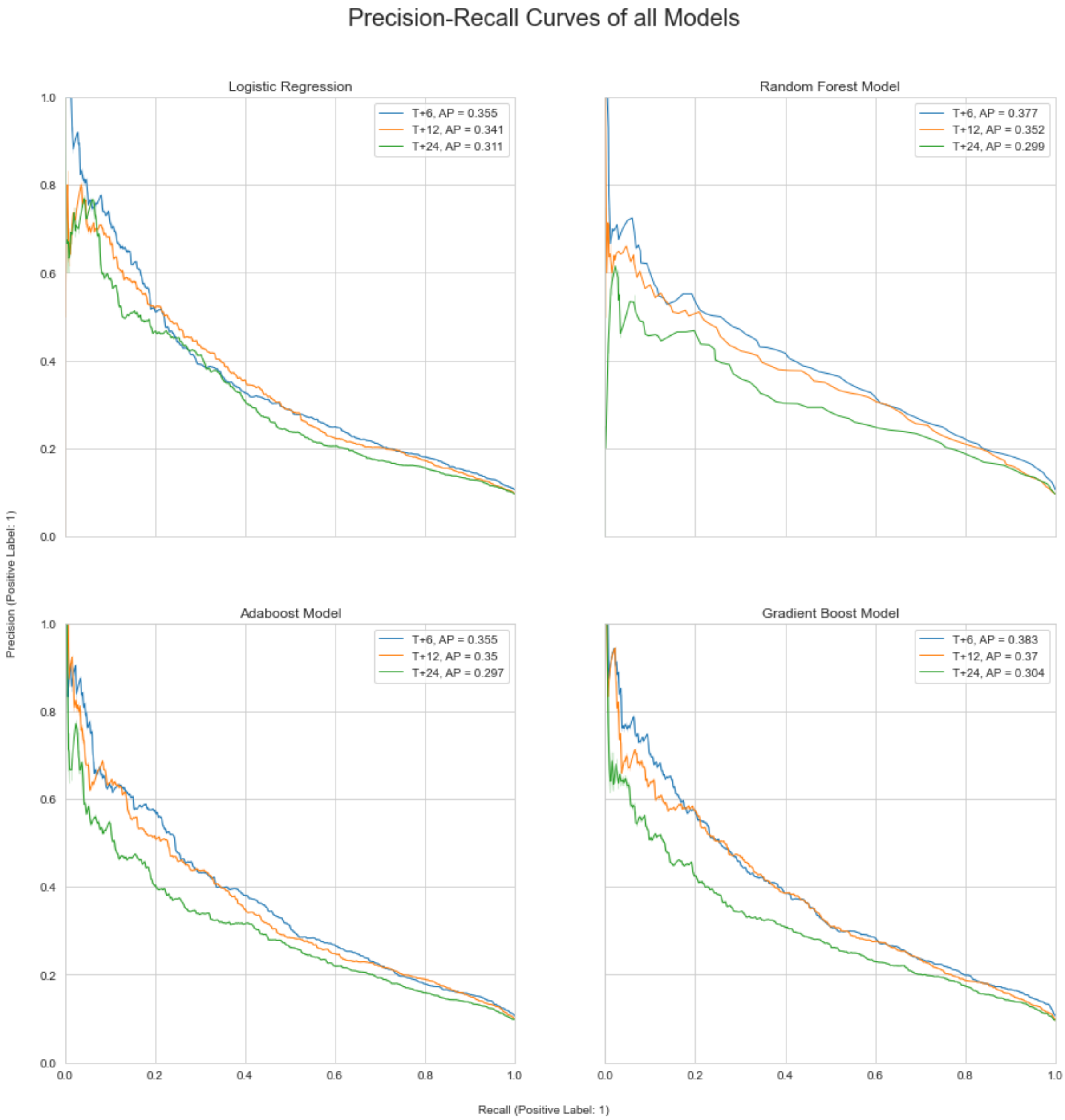

Figure 2: Precision-Recall curve

Appendix IX: Confusion matrices

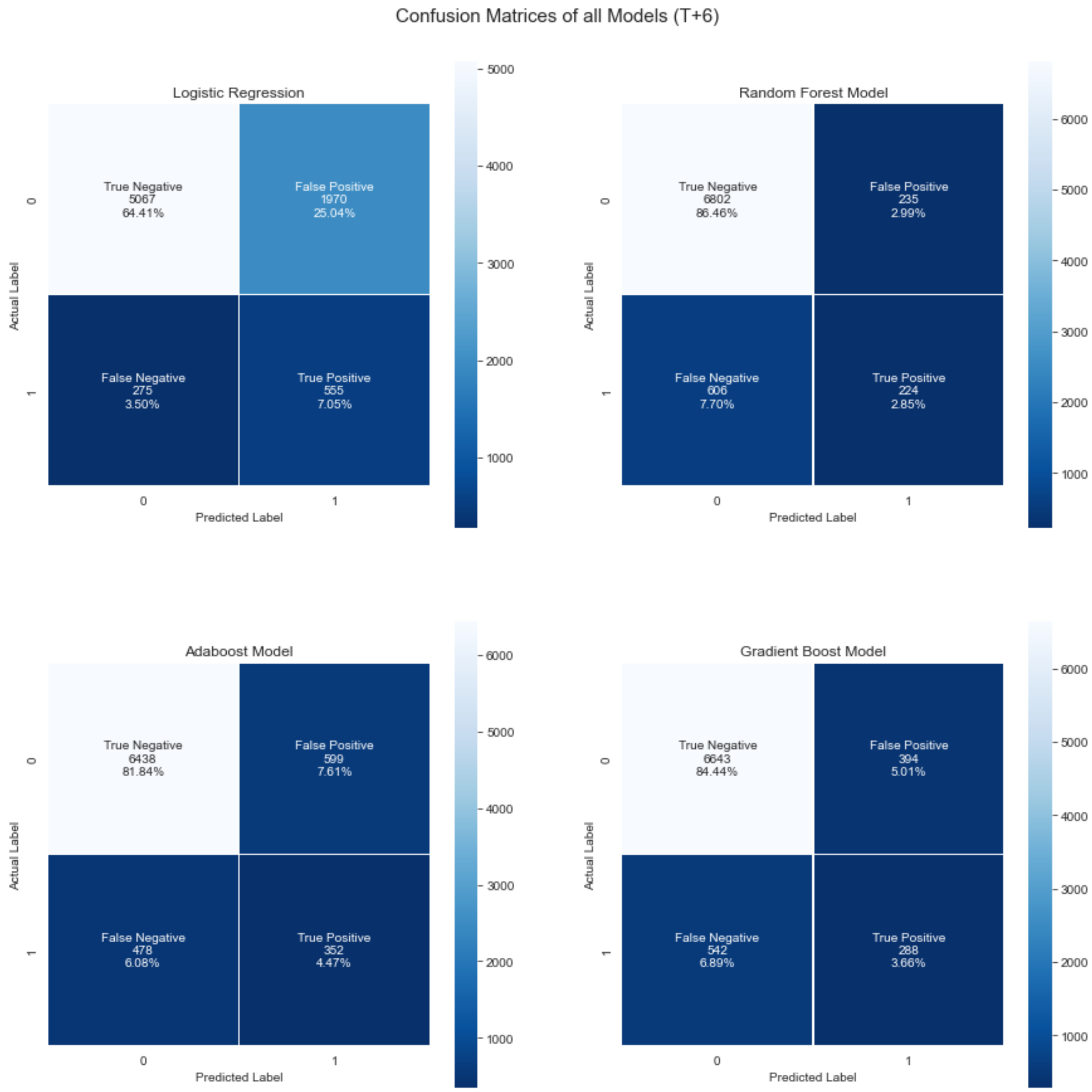

Figure 3: Confusion matrix (T+6)

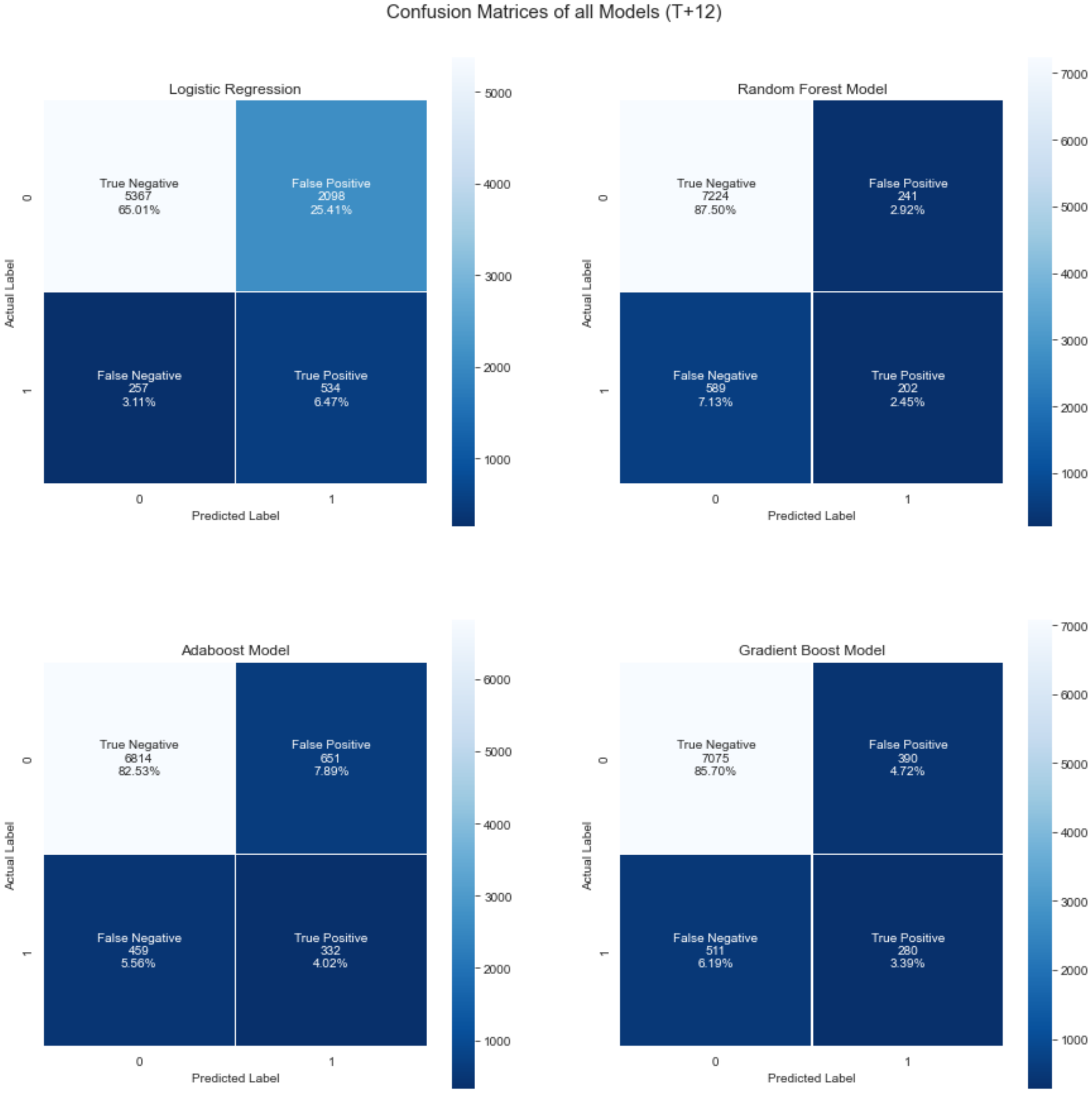

Figure 4: Confusion matrix (T+12)

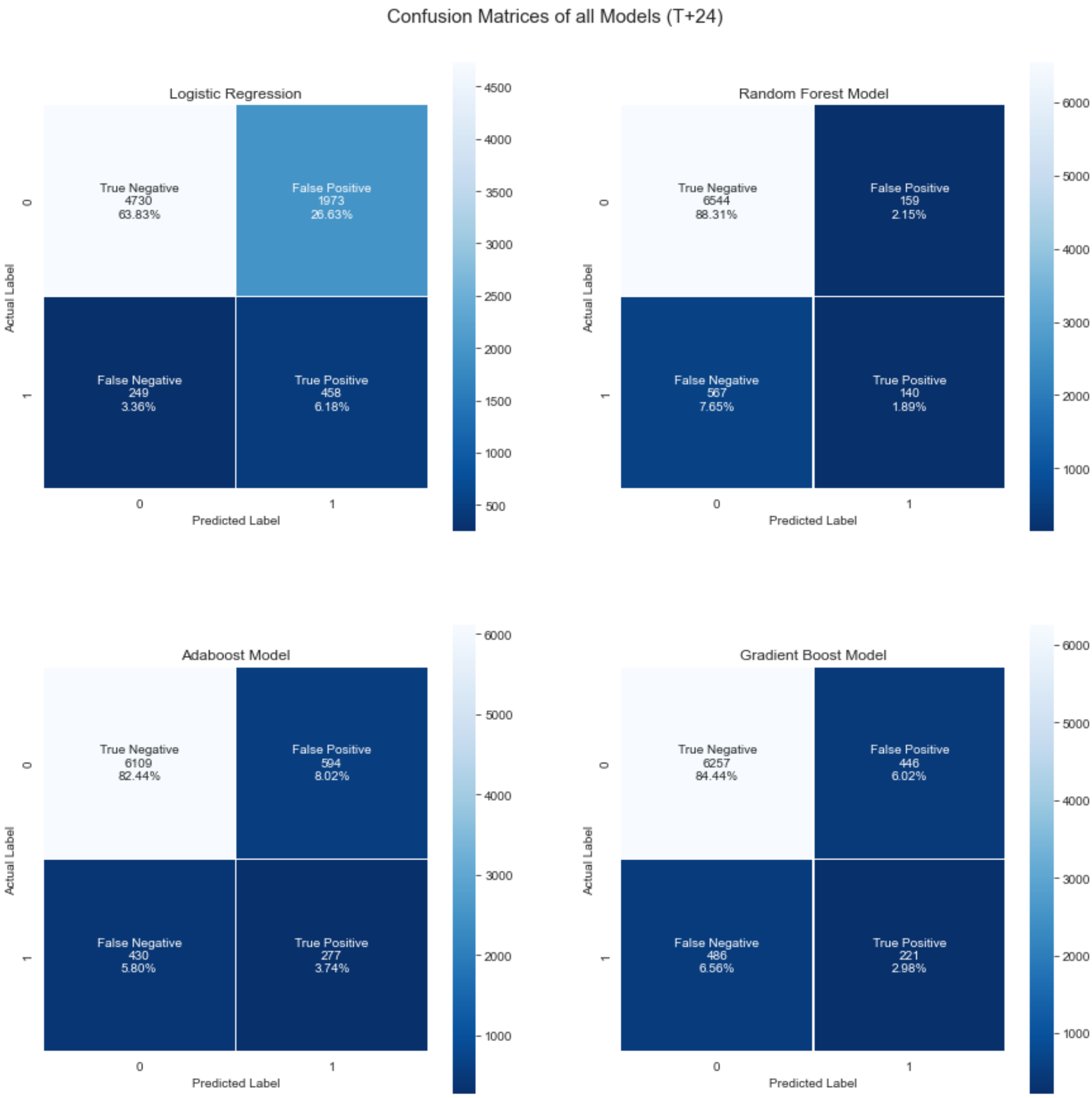

Figure 5: Confusion matrix (T+24)

Appendix X: Source code

See attached Jupyter notebook “HMA_Group5_Appendix10”.