PRESCRIPTIVE ANALYTICS CASE STUDIES ANALYSES

This document consists of two sections that contain the following information:

1. Analyses of prescriptive analytics case studies.
2. Analysis of the results produced by the case study analysis.

1. Case Study Analyses

Case Study 1: Case studies of clinical decision-making through prescriptive models based on machine learning [1]

Initiation

Pro requirement 1.1: The gaol of the analytics project is 'development of prescriptive models to support decision-making in clinical settings'.
Gen requirement 1.1: The goal of the analytics project must be defined.

Pro requirement 1.2: The analytics project must consider 'decision-making focused on treatment, follow-up and prevention of diseases'.
Gen requirement 1.2: The object to which the analytics is directed must be defined.

Acquisition

Pro requirement 1.3: The sources of the data that can be utilized in this project is 'public repositories'.
Gen requirement 1.3: The sources of data must be defined for the analytics project.

Pro requirement 1.4: Within the analytics project 'synthetic minority oversampling technique (SMOTE) to balance the classes before feeding the predictive and prescriptive algorithms'.
Gen requirement 1.4: The analytics project must define preprocessing done to the data before being inputted into the predictive model.

Pro requirement 1.5: The analytics project must see that the 'dataset was divided into 80% for training and validation and 20% for testing'.
Gen requirement 1.5: The splitting of the data set must be defined for the analytics project.

Analysis

Pro requirement 1.6: The analytics project predictive model(s) must be 'generated using a data-driven PSO-FCM approach'.
Gen requirement 1.6: The predictive model(s)/ algorithm(s) that must be used within the analytics project must be defined.

Pro requirement 1.7: The predictive analytics model must 'predict INR values' for case 1, 'predict mortal- ity from SD' for case 2.
Gen requirement 1.7: The expected output of the predictive analytics model must be defined for the analytics project.

Pro requirement 1.8: This analytics project must 'use mean absolute error (MAE), mean squared error (MSE), root mean squared error (RMSE) and $R2$ metrics' to evaluate the performance of the prescriptive analytics models.
Gen requirement 1.8: The performance measures for the created predictive model must be defined the analytics case.

Pro requirement 1.9: This analytics project must use ‘Prescriptive-FCM to generate prescriptive models’.
Gen requirement 1.9: The method/model(s) used to create the prescriptive analytics model or algorithm must be defined.
Pro requirement 1.10: The analytics project *R*2 must be used ‘as a measure of agreement between the actual data and that prescribed by the generated model’.
Gen requirement 1.10: The performance measures for the created prescriptive model/algorithm must be defined the analytics case.

Pro requirement 1.11: The analytics project must utilize ‘10-fold cross- validation technique to find the best hyperparameters in each model’.
Gen requirement 1.11: The method used to define hyperparameters for the analytics model(s)/algorithm(s) must be defined.

Visualisation
Gen requirement 1.12: The prescriptive model must generate the following for:
case study 1: ‘warfarin dose’,
case study 2: ‘prescribing treatment actions aimed at preventing patient death by SD’,
case study 3: ‘prescribe two crucial actions in the prevention of geohelminthiasis’.
Gen requirement 1.12: The output of the prescriptive analytics model must be defined, alongside the excepted utility of the models’ outputs.

Case study 2: Study on facility location of air taxi skyports using a prescriptive analytics approach [2]
Initiation
Pro requirement 1.1: The goal of the analytics project is ‘developing tools to gain insights into strategic decision planning for the optimal location of air taxi stations in any metropolitan city’.

Pro requirement 1.2: The analytics project must consider ‘infrastructure locations for air taxi operations in New York City (NYC)’.

Acquisition:
Pro requirement 1.3: The analytics project must utilize ‘air taxi NYC demand data’ and ‘input parameters required for the mathematical model’ (average rental cost, trips per day per 1000 population, average salary, and road facilities) which are obtained from various official sources (Department of City Planning, NYU Furman Center, and United States Census Bureau)’ within the analytics project.

Pro requirement 1.4: Not defined.

Pro requirement 1.5: Not defined.

Analysis
Pro requirement 1.6: Not defined.

Pro requirement 1.7: Not defined.

Pro requirement 1.8: Not defined.

Pro requirement 1.9: The analytics project must identify ‘ideal operating sites for air taxi services using the Clustering large applications (CLARA) technique’ and ‘integer programming model was developed to recommend specific stations to be built in multiple stages’.

Pro requirement 1.10: The prescriptive models must be ‘evaluate the effectiveness of the CLARA technique, the Davies–Bouldin index (DBI) metric values’.

Pro requirement 1.11: Not defined.

Pro requirement 2.1: Within the analytics project ‘Python programming language is used for developing the CLARA technique’ and ‘Excel solver was used to obtain the results of the phased approach’.
Gen requirement 2.1: The tools/software used within the project must be defined.

Presentation
Pro requirement 1.12: The prescriptive model must provide ‘Fourteen ideal operating sites for air taxi services’, and ‘recommend specific stations to be built in multiple stages’.

Case study 3: Optimizing outpatient appointment system using machine learning algorithms and scheduling rules: A prescriptive analytics framework [3]
Initiation
Pro requirement 1.1: The analytics project must develop a ‘prescriptive analytics framework to improve the performance of an AS with respect to patient satisfaction (measured using average patient waiting time and num- ber of patients unable to get an appointment for the day under consideration) and resource utilization (measured using average resource idle time, overflow time and overtime).’

Pro requirement 1.2: The analytics project must consider ‘patient scheduling’ in relation to ‘patient-specific no-show risk’.

Acquisition
Pro requirement 1.3: The analytics project must utilize ‘historical data is obtained from the scheduling system of a Family Medicine Clinic in Central Pennsylvania’.

Pro requirement 1.4: Not defined.

Pro requirement 1.5: For the defined data analytics project must ‘two-thirds of the data (50,858 patient visits) is used for training the ML algorithms, and the remaining one-third (25,427) is used for testing and evaluation’.

Analysis
Pro requirement 1.6: The analytics project must ‘consider five different supervised machine learning classifiers, namely: (i) Logistic Regression (Lin- ear Classifier) (ii) Neural Networks (Non-linear Classifier) (iii) Ran- dom Forests (Ensemble Classifier) (iv) Gradient Boosting (Ensemble Classifier) (v) Stacking (Ensemble Classifier)’.

Pro requirement 1.7: The prescriptive analytics project must ‘provides the probability value for the patient risk type, which is then converted to a high-risk or low-risk patient using a threshold parameter’.

Pro requirement 1.8: Within the analytics project the machine learning algorithms must be evaluated based on the ‘AUC values’.

Pro requirement 1.9: The analytics project must use a ‘simulation model ~~is developed~~ to evaluate each of the 11 scheduling rules (3 benchmark rules + 8 proposed rules)’.

Pro requirement 1.10: The analytics project ‘rules, are evaluated by two approaches : (1) using relative costs associated with the performance metrics; (2) using MCDM method to get relative weights associated with the performance metrics.’.

Pro requirement 1.11: The analytics project ‘parameters of the machine learning algorithms are tuned (optimized) using grid search with cross-validation’.

Pro requirement 2.1: The analytics project must use ‘R statistical computing software ~~is used~~ to develop (train and test) the ML algorithms’ and ‘Matlab’ to create the simulation model.

Presentation
Gen requirement 1.12: The analytics model must ‘make sequencing and overbooking decisions to generate the sched- ule’.

Case Study 4: Prescriptive analytics in public-sector decision-making: A framework and insights from charging infrastructure planning [4]
Initiation
Pro requirement 1.1: The analytics project must develop a ‘spatial analytics application that supports strategic decision-making regarding investments into urban charging infrastructure for electric vehicles’.

Pro requirement 1.2: The analytics project must be focus on ‘applying prescriptive analytics to the placement of charge points in urban areas’.

Acquisition
Pro requirement 1.3: The analytics project must utilize ‘Data on the usage of each CP was openly accessible in real time through an API’, ‘data on more than 61,000 POIs within the city centre of Amsterdam using the Google Places API’, and ‘data provided by the Dutch Cen- tral Bureau of Statistics (http://www.cbs.nl/) as control variables’.

Pro requirement 1.4: The analytics project must consist of ‘transforming the CP readings (a bi- nary variable indicating whether a given CP is in use) into charg- ing sessions’.

Pro requirement 1.5: Not defined.

Analysis
Pro requirement 1.6: The analytics project must create a predictive model using ‘beta regression model’.

Pro requirement 1.7: The predictive model must ‘pre- dict the utilization of a new CP in any given area’.

Pro requirement 1.8: The analytics project must ‘assess the out-of-sample predictive performance of the model, we conduct 10-fold cross validation (CV) and Monte Carlo cross validation (MCCV, also known as repeated random subsampling) (Batur, Bekki & Chen, 2018; Shao, 1993) using root-mean-square error (RMSE) compared to the naïve model (intercept only) as the performance criterion’.

Pro requirement 1.9: The analytics project must develop a prescriptive model using ‘a greedy heuristic and auxiliary variables to prescribe CP locations’.

Pro requirement 1.10: The results of the prescriptive analytics project must be evaluated based on ‘Average utilization over all CPs Distance to nearest CP: Median, Average, Maximum’.

Pro requirement 1.11: Not defined.

Pro requirement 2.1: Not defined.

Presentation
Pro requirement 1.12: The analytics model must provide ‘the placement of charge points for electric vehicles’.

Case study 5: Design and Evaluation of a Process-aware Recommender System based on Prescriptive Analytics [5]
Initiation
Pro requirement 1.1: The analytics project must ‘design of a prescriptive-analytics technique that recommends which actions/activities to perform next to optimize a certain KPI (Key Performance Indicator) of interest’.

Pro requirement 1.2: The analytics must be focused on ‘percentage of customers finding a new job’.

Acquisition
Pro requirement 1.3: The analytics project must utilize ‘real-life event data from a Dutch reintegration com- pany’.

Pro requirement 1.4: Not defined.

Pro requirement 1.5: Within the analytics project ‘event log L consisting of 12296 complete traces’, L must be ‘divided in three parts:
• log LTraining with 6148 traces (50%) was used to train the predictive-analytics oracle P;
• log LTesting with 3074 traces was used to evaluate oracle P, and to simulate the running traces, for which recommendations are created;
• log LSimilarity with 3074 traces was used to evaluate the quality of the recommendations.’

Analysis
Pro requirement 1.6: The analytics project must consider ‘three machine-learning techniques: Ran- dom Forest, Support Vector Machine, and Decision Tree’ when developing the predictive model.

Pro requirement 1.7: The predictive model created within the analytics project must predict 'KPI that is used in the evaluation is: the customer found a new job before reaching the maximum service duration'.

Pro requirement 1.8: The predictive models must be 'evaluated using two standard metrics: Accuracy, and AUC Score'.

Pro requirement 1.9: The analytics project must develop a 'prescriptive-analytics technique that recommends which actions/activities to perform next to optimize a certain KPI (Key Performance Indicator) of interest' which is then used on the defined data.

Pro requirement 1.10: The prescriptive model must be evaluated using 'Z Test to compare the average KPI value of the traces in $L^{G}_{\sigma_{r^{i}un}}$ with the average of the traces in $L^{B}_{\sigma_{r^{i}un}}$'.

Pro requirement 1.11: The hyper parameters must first be set to 'default parameters of the respective implementations in the SciPy package for Python' and then using 'hyper-parameter optimization'.

Pro requirement 2.1: The models created in the project must be 'imple- mented in Python'.

Visualisation
Pro requirement 1.12: The prescriptive analytics model must 'support the company's employee to choose the right interventions' by providing 'personalized recommendations'.

2. Analysis of Case Study Analyses

Shown below are all the generic requirements that were defined through the analysis of case studies.

Initiation
Gen requirement 1.1: The goal of the analytics project must be defined.

Gen requirement 1.2: The object to which the analytics is directed must be defined.

Acquisition
Gen requirement 1.3: The source(s) of data must be defined for the analytics project.

Gen requirement 1.4: The analytics project must define preprocessing done to the data before being inputted into the predictive model.

Gen requirement 1.5: The splitting of the data set must be defined for the analytics project.

Analysis
Gen requirement 1.6: The predictive model(s)/ algorithm(s) that must be used within the analytics project must be defined.

Gen requirement 1.7: The expected output of the predictive analytics model must be defined for the analytics project.

Gen requirement 1.8: The performance measures for the created predictive model/algorithm must be defined the analytics case.

Gen requirement 1.9: The method/model(s) used to create the prescriptive analytics model or algorithm must be defined.

Gen requirement 1.10: The performance measures for the created prescriptive model/algorithm must be defined for the analytics project.

Gen requirement 1.11: The method used to define hyperparameters for the predictive analytics model(s)/algorithm(s) must be defined.

Gen requirement 2.1: The tools/software used within the project must be defined.

Presentation
Gen requirement 1.12: The output of the prescriptive analytics model must be defined, alongside the excepted utility of the models' outputs.

Table 1 shows the definition and validation of generic requirements (case study analysis is shown in the appendix) for prescriptive analytics projects. In the table Y stands for Yes meaning that the generic requirement could be defined for a CS (Case Study) and N means the generic requirement could not be defined for the CS.

Gen(Prs)	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	1.10	1.11	1.12	2.1
CS1	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	
CS2	Y	Y	Y	N	N	N	N	N	Y	Y	N	Y	Y
CS3	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y
CS4	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	N	Y	N
CS5	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y

Table 1: Validation of generic requirements for prescriptive analytics.

The results of the validation show that Gen requirements 1.4 was not validated, meaning that the all the numeric labels for requirements proceeding Gen requirement 1.4 will be decremented by one accordingly. Meaning the labels for gen requirements will be changed from 1.5 to 1.4, 1.6 to 1.5, 1.7 to 1.6, 1.8 to 1.7, 1.9 to 1.8, 1.10 to 1.9, 1.11 to 1.10, and 1.12 to 1.11.

The finalized generic requirements templates is as follows:

Initiation
Gen requirement 1.1: The goal of the analytics project must be defined.

Gen requirement 1.2: The object to which the analytics is directed must be defined.

Acquisition
Gen requirement 1.3: The sources of data must be defined for the analytics project.

Gen requirement 1.4: The splitting of the data set must be defined for the analytics project.

Analysis
Gen requirement 1.5: The predictive model(s)/ algorithm(s) that must be used within the analytics project must be defined.

Gen requirement 1.6: The expected output of the predictive analytics model must be defined for the analytics project.

Gen requirement 1.7: The performance measures for the created predictive model/algorithm must be defined the analytics case.

Gen requirement 1.8: The method/model(s) used to create the prescriptive analytics model or algorithm must be defined.

Gen requirement 1.9: The performance measures for the created prescriptive model/algorithm must be defined for the analytics project.

Gen requirement 1.10: The method used to define hyperparameters for the predictive analytics model(s)/algorithm(s) must be defined.

Gen requirement 2.1: The tools/software used within the project must be defined.

Presentation
Gen requirement 1.11: The output of the prescriptive analytics model must be defined, alongside the excepted utility of the models' outputs.

References

[1] W. Hoyos, J. Aguilar, M. Raciny, and M. Toro, “Case studies of clinical decision-making through prescriptive models based on machine learning,” Computer Methods and Programs in Biomedicine, vol. 242, p. 107829, Dec. 2023, doi: 10.1016/J.CMPB.2023.107829.

[2] A. Amitanand Sinha and S. Rajendran, “Study on facility location of air taxi skyports using a prescriptive analytics approach,” Transportation Research Interdisciplinary Perspectives, vol. 18, p. 100761, Mar. 2023, doi: 10.1016/J.TRIP.2023.100761.

[3] S. Srinivas and A. R. Ravindran, “Optimizing outpatient appointment system using machine learning algorithms and scheduling rules: A prescriptive analytics framework,” Expert Systems with Applications, vol. 102, pp. 245–261, Jul. 2018, doi: 10.1016/J.ESWA.2018.02.022.

[4] T. Brandt, S. Wagner, and D. Neumann, “Prescriptive analytics in public-sector decision-making: A framework and insights from charging infrastructure planning,” European Journal of Operational Research, vol. 291, no. 1, pp. 379–393, May 2021, doi: 10.1016/J.EJOR.2020.09.034.

[5] M. d. Leoni, M. Dees and L. Reulink, "Design and Evaluation of a Process-aware Recommender System based on Prescriptive Analytics," 2020 2nd International Conference on Process Mining (ICPM), Padua, Italy, 2020, pp. 9-16, doi: 10.1109/ICPM49681.2020.00013.