CLASSIFICATION OF STUDENT'S AIR TRAFFIC CONTROL SKILL USING LOGISTIC REGRESSION

[1]Liber Tommy Hutabarat, [2]Hairul Amren S, [3]Marzuki Sinambela, [4]Tonni Limbong

[1,2] Akademi Teknik dan Keselamatan Penerbangan Medan
[3]Departement of Physics, FMIPA, Universitas Sumatera Utara, Medan, Indonesia
[4]Computer Science Faculty, Chatolic University of Saint Thomas North Sumatera, Medan, Indonesia
email: limasodara@gmail.com, hairul.amren@yahoo.com, sinambela.m@gmail.com,
tonni.budidarma@gmail.com

ABSTRACT

The classification of student's air traffic control skills at Akademi Teknik dan Keselamatan Penerbangan Medan (ATKP) is very interesting to evaluate and look at the performance of the student. In this study, we compute the student's air traffic control (ATC) skill data to classify and evaluate the model and performance of the dataset. The computation of the dataset using the logistic regression approach based on Sk-learn by training and test data. The data was collected from ATKP for twenty samples. The result of this study indicates the logistic regression classifier is the best algorithm for this classification problem, offering good values in terms of accuracy, true negative rate, and true positive rate.

Keywords: Classification, ATC, Logistic Regression.

INTRODUCTION

The aviation industry has faced incredible challenges over the last few years (Bisignani, 2011; Duncan & Bragadish, 2017; KPMG, 2018), such as the emphasis on aviation and commercial safety. The ACT program began with the ATKP institution as part of the Department of Transportation of Indonesia. The primary purpose of the ACT programs was to "provide the skill with qualified applicants possessing a broad-based knowledge of the aviation industry to fill ATC assign. The institutions in the ATKP Medan, Indonesia which offer associate, and bachelor degrees. ACT specialist is tasked with promoting safe, orderly, and expeditious flow of air traffic. In general, the ACT specialist job consists of complex tasks demanding a high degree of skill (Office & Management, 1988;

OPM, 2009) and active application of unique cognitive abilities such as information processing, reasoning and decision making. According to a report filed by ATKP Medan, applicants for ATC must pass an aptitude test and meet certain age, medical, and security qualifications. In this study, we focused to evaluate the model of the students for ATC skill to investigate and got of decision making for student's skill by using the logistic regression based on sk-learn by training and test the data.

RELATED WORK

This study shows a relationship between an individual's cognitive learning style and ability to do well on simulated ATC practical exercise. The individual/student learning style is an important factor in determining an individual's ability to adapt

to a particular learning challenge and discussion making (Bucos & Drăgulescu, 2018). Some researcher tries to describe in great detail the selection, training, and other labor-related issues (Galotti, 1991). Other (Coyne, Rice, Winter, Tamilselvan, & Drechsel, 2017; Pierce, Broach, Bleckley, & Byrne, 2013), describe the emotional aspects of the career into ATC facilities on busy days and trainees and controllers.

DATA AND METHODS

The sample data of student's ATC skills were collected in ATKP Medan, Indonesia. To be included in this sample, the candidate must have passed the test during the application process and must have completed the ability test.

Table 1. The Student's ATC skill report

Name	Overall Grade	Obedient	Research	Project Score	Recommend
A1001	B	Y	72	72	Yes
A1002	B	Y	69	90	Yes
A1003	A	N	90	95	Yes
A1004	C	N	47	57	No
A1005	B	Y	76	78	Yes
A1006	A	N	71	83	Yes
A1007	A	Y	88	95	Yes
A1008	E	Y	40	43	No
A1009	C	Y	64	64	Yes
A1010	D	Y	38	60	No
A1011	D	N	58	54	No
A1012	D	N	40	52	No
A1013	C	Y	65	81	Yes
A1014	B	N	78	72	Yes
A1015	C	Y	50	53	Yes
A1016	B	Y	69	75	Yes
A1017	A	Y	88	89	Yes
A1018	E	Y	18	32	No
A1019	D	N	46	42	No

The methods of this study using logistic regression. This study focused to classify and evaluate the data skill reports of ATC students (Davis & Offord, 2013; Pierce et al., 2013).

Abstractly, regression is the problem of calculating a conditional expectation

$$E[Y|X = x]$$

The form occupied by this expectation is different depending on the assumptions of how the data were generated, the first, assuming (Y|X=x) to be normally distributed yields with classical linear regression. Second, assuming a Poisson distribution yields Poisson regression, and the last assuming a Bernoulli distribution yields logistic regression (Hall, 2000; Zaninotto & Falaschetti, 2011). The term "regression" has also been used more generally than this, including approaches like quantile regression, which estimates a given quantile of $(Y|X=x)$.

RESULT AND DISCUSSION

The result in this study, compute and produced from preparing the student data by feature extraction to be overall grade, obedient, research score, project score and recommend.

	OverallGrade	Obedient	ResearchScore	ProjectScore
0	B	Y	0.551362	0.231711
1	B	Y	0.395006	1.210044
2	A	N	1.489500	1.481803
3	C	N	-0.751608	-0.583567
4	B	Y	0.759837	0.557822
5	A	N	0.499243	0.829581
6	A	Y	1.385263	1.481803
7	E	Y	-1.116439	-1.344493
8	C	Y	0.134412	-0.203104
9	D	Y	-1.220677	-0.420512
10	D	N	-0.178301	-0.746623
11	D	N	-1.116439	-0.855327
12	C	Y	0.186530	0.720877
13	B	N	0.864075	0.231711
14	C	Y	-0.595252	-0.800975
15	B	Y	0.395006	0.394766
16	A	Y	1.385263	1.155692
17	E	Y	-2.263053	-1.942364
18	D	N	-0.803727	-1.398845

Figure 1. Numerical feature training set

The classify the data by simple making in labels. The numerical training data in this result show in Fig. 1.

The model numerical shown in fig. 2. The training model in numerical student data set shown the true negative and positive both of in research score and project score.

	ResearchScore	ProjectScore	...	Obedient_N	Obedient_Y
0	0.551362	0.231711	...	0	1
1	0.395006	1.210044	...	0	1
2	1.489500	1.481803	...	1	0
3	−0.751608	−0.583567	...	1	0
4	0.759837	0.557822	...	0	1
5	0.499243	0.829581	...	1	0
6	1.385263	1.481803	...	0	1
7	−1.116439	−1.344493	...	0	1
8	0.134412	−0.203104	...	0	1
9	−1.220677	−0.420512	...	0	1
10	−0.178301	−0.746623	...	1	0
11	−1.116439	−0.855327	...	1	0
12	0.186530	0.720877	...	0	1
13	0.864075	0.231711	...	1	0
14	−0.595252	−0.800975	...	0	1
15	0.395006	0.394766	...	0	1
16	1.385263	1.155692	...	0	1
17	−2.263053	−1.942364	...	0	1
18	−0.803727	−1.398845	...	1	0

Figure 2. Training feature numerical

The logistic regression modeling can show in Fig. 3. The modeling in this study has a model parameters for recommended variable.

```
LogisticRegression(C=1.0, class_weight=None, dual=False, fit_intercept=True,
        intercept_scaling=1, max_iter=100, multi_class='warn',
        n_jobs=None, penalty='l2', random_state=None, solver='warn',
        tol=0.0001, verbose=0, warm_start=False)
```

Figure 3. The model parameter based on logistic regression

The result of this study, based on the logistic regression indicated the accuracy 100 %. And the classification statistic show in fig. 4. The performance of evaluation model in this case using sk-learn.

```
('Accuracy:', 100.0, '%')
Classification Stats:
              precision   recall   f1-score   support

         No      1.00      1.00      1.00         7
        Yes      1.00      1.00      1.00        12

  micro avg      1.00      1.00      1.00        19
  macro avg      1.00      1.00      1.00        19
weighted avg     1.00      1.00      1.00        19
```

Figure 4. Performance of evaluation model

The simple evaluation model on training data focused to predict the labels and actual labels in recommend parameter.

CONCLUSSION

This paper describes a study in classifying and evaluates the student's ATC performance by using only report data generated twenty samples from ATKP Medan, Indonesia. The experimental result shows that the student can be identified in an early stage event from performance skill data. The result showed that Logistic Regression Classifier is the best algorithm for this classification problem, offering good values in terms of accuracy, true negative rate, and true positive rate.

REFERENCE

Bisignani, G. (2011). IATA Vision 2050, (February), 62–63.

Bucos, M., & Drăgulescu, B. (2018). Predicting student success using data generated in traditional educational environments. *TEM Journal*, *7*(3), 617–625. https://doi.org/10.18421/TEM73-19

Coyne, W., Rice, S. C., Winter, S., Tamilselvan, G., & Drechsel, P. V. (2017). Simulation challenges - Student perception of air traffic control simulation. *International Journal of Aviation, Aeronautics, and Aerospace*, *4*(3). https://doi.org/10.15394/ijaaa.2017.1088

Davis, L. J., & Offord, K. P. (2013). Logistic regression. *Emerging Issues and Methods in Personality Assessment*, 273–283. https://doi.org/10.4324/9780203774618-23

Duncan, A., & Bragadish, N. (2017). The Future of Air Travel: Eight Disruptive Waves of Change. *Cognizant*, (June). Retrieved from https://www.cognizant.com/whitepapers/the-future-of-air-travel-eight-disruptive-waves-of-change-codex2566.pdf

Galotti, V. P. (1991). Cognitive Learning Styles And Their Relation To Air Traffic by.

Hall, D. B. (2000). Zero-inflated poisson and binomial regression with random effects: A case study. *Biometrics*, *56*(4), 1030–1039. https://doi.org/10.1111/j.0006-341X.2000.01030.x

KPMG. (2018). The Aviation Industry Leaders Report 2018: Navigating the Cycle. Retrieved from *www.airlineeconomics.co*

Office, U. S., & Management, P. (1988). *Position Classification Standard for Occupational Therapist Series* , (April), 1–17.

OPM. (2009). Introduction to the Position Classification Standards, (August 1991), 1–79. Retrieved from *http://www.opm.gov/policy-data-oversight/classification-qualifications/classifying-general-schedule-positions/positionclassificationintro.pdf*

Pierce, L. G., Broach, D., Bleckley, M. K., & Byrne, C. L. (2013). Using biodata to select air traffic controllers. *Proceedings of the Human Factors and Ergonomics Society*, (October), 823–827. https://doi.org/10.1177/1541931213571179

Zaninotto, P., & Falaschetti, E. (2011). Comparison of methods for modelling a count outcome with excess zeros: Application to Activities of Daily Living (ADL-s). *Journal of Epidemiology and Community Health*, *65*(3), 205–210. https://doi.org/10.1136/jech.2008.079640