International Journal Of Artificial Intelligence Research
Vol 5, No 1, June 2021, pp. 53-64
ISSN: 2579-7298

Uplift modeling VS conventional predictive model: A reliable machine learning model to solve employee turnover

Davin Wijaya, Jumri Habbeyb DS, Samuelta Barus, Beriman Pasaribu, Loredana Ioana Sirbu, Abdi Dharma
[a] *Researcher in the Computer Science department Universitas Prima Indonesia, Indonesia*
[b] *Researcher in the Department of Language, Literature and Intercultural Studies The University of Florence Italy*
[1] *wijayadavin@gmail.com**
* *corresponding author*

ARTICLE INFO

Article history:
Received 06-05-2020
Revised 21-10-2020
Accepted 19-07-2021

Keywords:
Machine Learning
Uplift Modeling
Employee Turnover Prediction
Extreme Gradient Boosting
Lai's generalized weighed uplift

ABSTRACT

Employee turnover is the loss of talent in the workforce that can be costly for a company. Uplift modeling is one of the prescriptive methods in machine learning models that not only predict an outcome but also prescribe a solution. Recent studies are focusing on the conventional predictive models to predict employee turnover rather than uplift modeling. In this research, we analyze whether the uplift model has better performance than the conventional predictive model in solving employee turnover. Performance comparison between the two methods was carried out by experimentation using two synthetic datasets and one real dataset. The results show that despite the conventional predictive model yields an average prediction accuracy of 83.35%; it only yields a success rate of 50% to target the right employee with a retention program on the three datasets. By contrast, the uplift model only yields an average accuracy of 70.03% but yields a consistent success rate of 100% in targeting the right employee with a retention program.

Copyright © 2020 International Journal of Artificial Intelligence Research.
All rights reserved.

I. Introduction

A. *Employee turnover*

Employee turnover — also identified as churn or attrition — is the voluntary loss of talent in the workforce that can be costly for an organization, as Peter et al [1] advocate the view of this field of study over the past 100 years and concludes that the cumulative knowledge may allow us to not only theorize about employee turnover but also test various kinds of prediction models about employee turnover. According to [2], an employee retention program may be an effective solution to solve employee turnover and further reducing turnover costs and significantly increase the company's productivity. Moreover, employee retention programs may apply various strategies to reduce employee turnover, for instance, employee training and development, promoting work-life balance, and of course, employee promotion. Furthermore, Katharina's [3] study suggests an evidence-based approach to making better decisions for *Human Resource* (HR) with analytics techniques and generally, there are three levels of analytics as visualized in Figure 1.

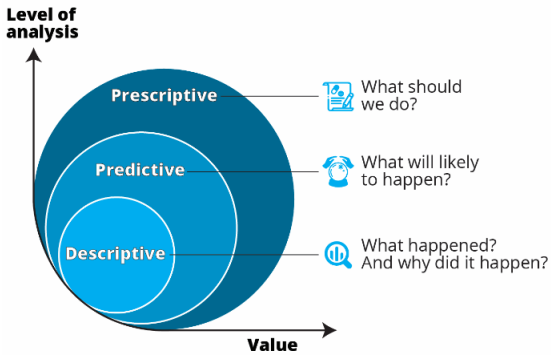

Figure 1. The three-level of analytics

B. Human resource analytics (HR Analytics)

Katharina further points out that human resources analytics also consist of three levels:

1) *Descriptive*: The first level of HR analytics is describing the relationships between current and historical data — e.g. what causes an employee to churn.
2) *Predictive*: The next level after descriptive analytics, is predictive. It uses current and historical data to predict future outcomes and gather meaningful information before being used for prescriptive analytics — e.g. predicting employee turnover.
3) *Prescriptive*: In contrast with predictive analytics, prescriptive analytics is on a higher level than predictive analytics because it is not only used to predict employee turnover but also decision options to optimize the workforce — e.g. treating employees with high turnover rate.

Furthermore, predictive or prescriptive analytics may apply machine learning approaches to predict and solve many kinds of problems that could happen in the future [4]. There are many applications of those approaches in the literature that not only applied to solve HR problems [5] but also applied widely to solve business problems [4], healthcare problems [6], industry problems [7], etc. In the application of predictive analytics for HR, a conventional binary classification model like *employee turnover prediction* (ETP) is frequently studied in the literature to predict employee turnover [8]–[15]. A survival analysis model [16] may also be applied as predictive analytics to predict employee with the highest turnover probability at a given time.

Nevertheless, we may upgrade the predictive analytics models into the prescriptive analytics model by focusing on the decision making capability of the predictive model to not only predict the future but also prescribe optimal decisions [17]. There are various prescriptive analytics models for HR that have been studied in the literature. For instance, an HR recruitment model [18]–[20] and *employee turnover uplift* (ETU) model [21], [22]. Furthermore, Floris et al [23] suggest that uplift modeling may be widely applied, for instance in marketing, personalized medicine, and political election. Floris et al additionally explain that uplift modeling may solve the feedback loop issue that occurred when applying prescriptive analytics with a conventional predictive model.

C. Contributions

Floris *et al*. [24] and Eva [25] conclude that uplift modeling has a better prescriptive analytics performance than the conventional predictive model to solve employee churn problems. But the same study has never been conducted for employee turnover problems, to the authors' best knowledge. Therefore, we conduct an experimental case study with two synthetic datasets and one real dataset. In summary, this research aims to provide empirical evidence in the merits of the *employee turnover prediction* model versus the *employee turnover uplift* model as prescriptive analytics to solve employee turnover with an experimental case study of three different datasets.

II. Literature

A. Conventional predictive model

In this paper, we discuss a conventional predictive model for HR data, which is the *employee turnover prediction* (ETP). In the application of the ETP model, the same binary classification method to predict customer churn may be applied for employee turnover problems. The target variable — also known as the dependent variable, or outcome variable — of the ETP model is the binary parameter of the employee turnover. Moreover, Mishra et al [26] explained a prescriptive method for the ETP model by targeting employees with high turnover probability, thus we will use that method to measure the prescriptive performance of the ETP model.

B. Extreme gradient boosting (XGB)

A systematic comparative study of turnover prediction by Zhao [15] concluding that the *extreme gradient boosting* (XGB) algorithm or also commonly known as XGBoost is reliable in predicting employee turnover. XGB is a tree-based ensemble machine learning algorithm that was introduced in 2014 by Chen [27]. In Contrast with the gradient boosting method, XGB uses a regularization term to decrease the overfitting impact. Furthermore, Nielsen [28] also report the impressive track record of XGB that 17 of 29 challenge winning solutions on Kaggle in 2015 were using XGB. Therefore, we will use XGB as our machine–learning algorithm to build the ETP model and ETU model.

C. Uplift modeling

Uplift modeling (also identified as incremental, true lift, or net modeling) uses the treatment and control group to measure the effectiveness of a treatment [29]. *Lai's generalized weighed uplift method (LGWUM)* is one of several strategies available for uplift modeling, as in [30]. LGWUM was introduced in 2002 by Lo and generalized later in 2014 by Kane *et al* [31]. This uplift modeling method uses a four-quadrant approach to measure the uplift score that will be used as a prescriptive result to target the right individuals with the measured treatment. Kane *et al* also report the merits of LGWUM in comparison with three different uplift strategies and three different datasets.

Therefore, we will use LGWUM as the uplift strategy to get the uplift score in this research. By using an employee retention program as a treatment, the treatment and control group can be identified:

- *Control Group (C):* Employees who were not treated with the retention program.
- *Treated Group (T):* Employees who were treated with the retention program.

Four target classes are generated by fitting the treatment status and employee turnover status as visualized in the left part of Figure 1:

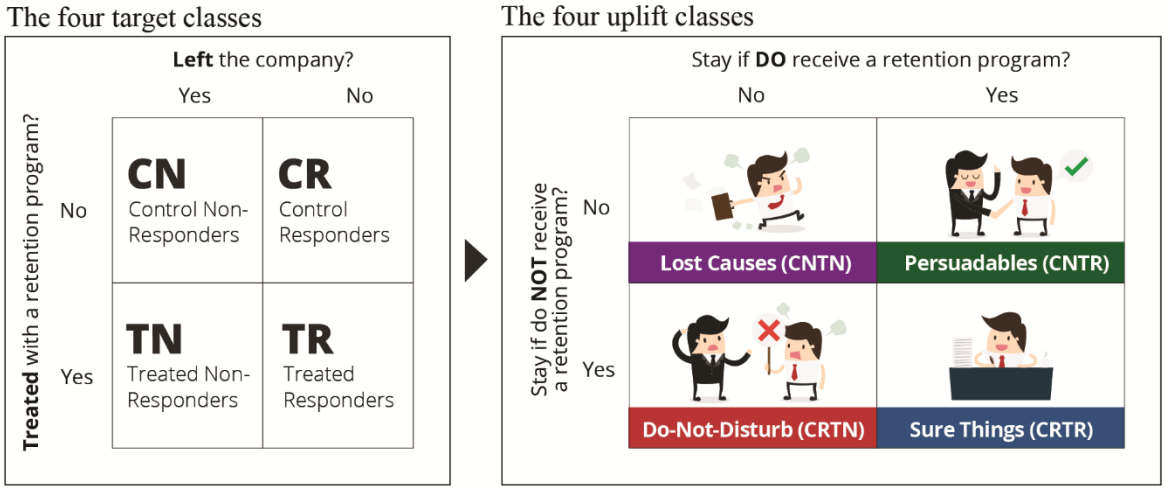

Figure 2. Target classes and uplift classes for employee turnover problems

- *Control Non–responders (CN):* Employees who have not been treated with the retention program and left. We want to find the persuadables in this group.
- *Control Responders (CR):* Employees who have not been treated with the retention program and stay. We want to avoid targeting this group because we do not need to treat them to make them stay, and there is a possibility that some of them are *Do-Not-Disturbs*.
- *Treated Non–Responders (TN):* Employees who have been treated with the retention program but left. We want to avoid treating this group because they will leave if treated, and there is a possibility that some of them are *Do-Not-Disturbs* that will stay if left untreated.
- *Treated Responders (TR):* Employees who have been treated with the retention program and stayed. We also want to find the persuadables in this group.

Therefore, by combining *CN* with *CR* and *TN* with *TR* will result in *C*, and *T*, *respectively*. Thus by fitting the treated and control group from those four target classes as visualized in the right table of Figure 2 will yield four possible theoretical uplift classes (also known as the four quadrants):

- *Do-Not-Disturb (CRTN):* Sometimes referred to as *sleeping dogs*, employees who will be driven away if treated.
- *Lost Causes (CNTN):* Employees who will leave whether treated or not.
- *Sure Things (CRTR):* Employees who will stay whether treated or not.
- *Persuadables (CNTR):* Employees who are willing to leave but will stay if treated. We want to target this group to reduce overall employee turnover in the company.

After predicting the four target classes that are visualized in Figure 2 with a machine–learning algorithm, four probability results are generated. Where *P* is the probability result, thus uplift score is calculated with LGWUM as:

$$Uplift\ Score = P\left(\frac{CN}{C}\right) + P\left(\frac{TR}{T}\right) - P\left(\frac{CR}{C}\right) - P\left(\frac{TN}{T}\right) \quad (1)$$

III. Human resource datasets

This section reports the description of the three public datasets that are included in this research:

1) *Dataset 1:* A synthetic HR dataset shared by Giri Pujar on Kaggle. This dataset consisting of 10 features with a sample of 14999 employees that previously studied by [8]–[11] using the ETP model — retrieved from https://www.kaggle.com/giripujar/hr-analytics.

2) *Dataset 2:* The *IBM Watson Analytics* synthetic dataset shared by Pavan Subhasht on Kaggle. This dataset consisting of 35 features with a sample of 1,470 employees, studied by [12]–[15] using the ETP model — retrieved from https://www.kaggle.com/pavansubhasht/ibm-hr-analytics-attrition-dataset.

3) *Dataset 3:* A real dataset shared by Eduard Babushkin. This dataset consisting of 16 features with a sample of 997 employees, this dataset has been studied for a survival analysis model — retrieved from https://edwvb.blogspot.com/2017/10/employee-turnover-how-to-predict-individual-risks-of-quitting.html.

The three datasets provided in this research have not been studied for uplift modeling. Moreover, no treatment — i.e. retention program — that has been identified for the three datasets.

IV. Methodology

In this section, information about the experimental methodology conducted in this research is provided. The methodology consists of four steps, which are data preprocessing, treatment identification, machine-learning modeling, and evaluation that are visualized in Figure 3 as the machine learning pipeline in this research. Furthermore, the experimental case study in this research is conducted with the Python programming language.

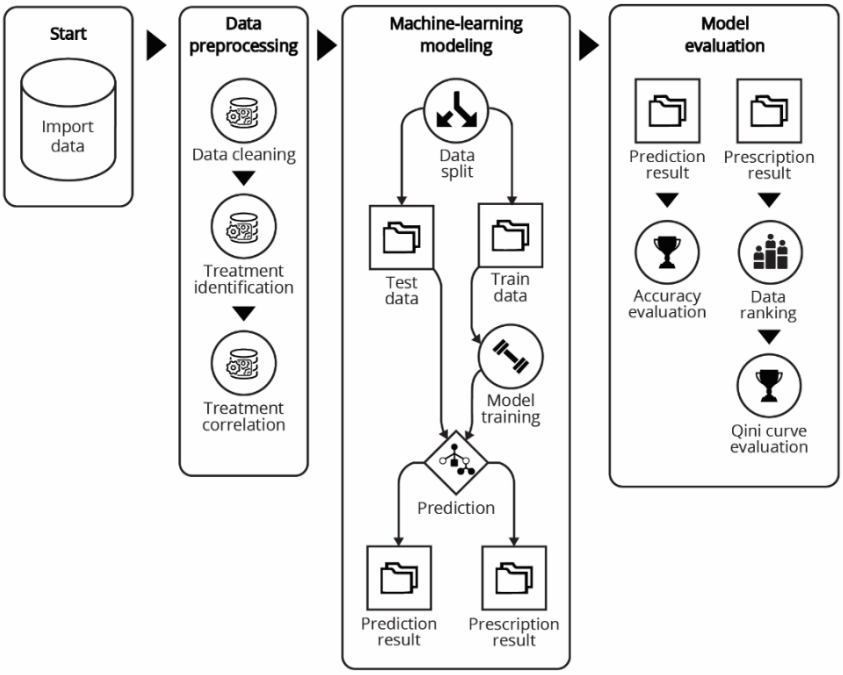

Figure 3. Machine learning pipeline

A. *Data preprocessing*

The first step in building the ETP model and ETU model in this research is data preprocessing. Data preprocessing is a technique to transform the data before feeding it into the algorithm. In this research, we conduct three steps of data preprocessing, which are data cleaning, treatment identification, and treatment correlation.

1) Data cleaning

This research uses one-hot encoding and label encoding using the Scikit–learn and Pandas package, respectively. Label encoding will be used to encode string data with hierarchy (e.g. Employee turnover, or education level) and one-hot encoding to encode categorical data (e.g. job position, or department). Previous works on the three datasets provided in this research applied various techniques to improve accuracy performance.

By contrast, this research only uses a treatment identification approach as a feature engineering method to unlock the prescriptive capabilities of the ETP model and ETU model. Thus, this research will not conduct sophisticated feature engineering to get more accuracy. The target variable for the ETP model will be the employee turnover status with a binary parameter. On the other hand, the ETU model will use the four target classes, *CN*, *CR*, *TN*, and *TR* with numerical values of 0, 1, 2, and 3, respectively.

2) Treatment Identification

The treatment identification approach in this research is proposed to unlock the prescriptive capability of a dataset with an unidentified treatment. The treatment identification criterion to specifically identify a feature that can be used as an employee retention program is formulated under these assumptions:

- *Actionable*: The treatment must be actionable, so it can be implemented by the company as an employee retention program — e.g. promotion or training.
- *Correlated with the target variable*: The treatment must have a significant treatment correlation with employee turnover to reduce employee turnover. We describe the treatment correlation as the strength of a linear association between the treatment and the target variable — i.e. employee turnover. Moreover, a significance level of $p > 0.005$ or $p > 0.05$ is generally used in the literature [32].
- *Control group availability*: There must be a group of employees who do not receive the treatment as a requirement to assess the treatment's performance and build the uplift models.

3) Treatment correlation

Note that by inverting the treatment parameter — from 0 to 1, and vice versa — will also inverting the treatment correlation value, and will give a different prescriptive scenario. Therefore, duplicating the three datasets and invert the treatment parameter will result in a total of six experiments in this research. Pearson correlation function from the Pandas package is used to calculate the treatment correlation before building the model.

B. Machine–learning modeling

This section will describe the machine–learning modeling methods in this research:

1) *Data split*: The data split process in this research divides the dataset into two data with, namely, training data and testing data with the proportion of 70% and 30%, respectively. Moreover, we apply random stratified sampling with the treatment variable to prevent data imbalance [24].

2) *Model training*: XGB algorithm is used to train the model with the training data. Note that hyper-tuning is not used in this research, therefore, we just apply the default setting provided by the XGB package.

3) *Prediction*: Testing data is used to predict the model that has been trained with the XGB algorithm, results in prediction results and probability results. For the ETP model, the probability result is used as a prescription result. By contrast, the ETU model uses the LGWUM to calculate the uplift score as the prescription result.

C. Evaluation method

This section shows the evaluation methods in this research:

1) *Accuracy evaluation*: In previous works, accuracy is the key measurement of the model's performance. Therefore, we measure the accuracy to assess the predictive performance of the ETP model and the ETU model. We use the Scikit–learn package to measure the accuracy score in this research.

2) *Qini curve*: Qini curve is used to evaluate the prescriptive performance of the ETP model and ETU model in this research. As demonstrated by Radcliffe [29], the Qini curve is a generalization of the Gini curve to measure uplift based on the calculation of every segment in the population:

α = proportion targeted with a treatment

N = total population of employees

With those definitions, uplift can be normalized into percentage as:

$$uplift\ (\alpha) = N\alpha\left(\frac{TR}{T} - \frac{CR}{C}\right) \quad (2)$$

Moreover, line plots from the Seaborn package are used to plot the Qini curve for the ETP model and ETU model, then a random model curve is added to the plot. Random model is a model that cannot differentiate the positive and negative uplift and results in a straight line from (0,0) to [1, uplift(N)]. After measured uplift in all proportion, the qini coefficient can be formulated as:

$$qini\ coefficient = \sum_{i=0}^{N-1} uplift\ -\ random\ model\ curve \quad (3)$$

To assess whether a model is successfully targeting the right employees — i.e. measuring the prescriptive performance —, we compare the model's Qini curve with the random model. Therefore, a successful model is the one that has a Qini coefficient greater than 5% ($P > 0.05$). To further analyze the prescriptive performance, the Swarm plot from the Seaborn package is used to visualize the target class distribution in this research that was observed to evaluate the target classes that tend to be recorded.

V. Results

This section reports the results of the experimental case study with the three datasets. After exploring and cleaning the three datasets we successfully identify a treatment for each dataset as presented in Table 1 together with the descriptions about the identified treatment.

Table 1. Treatments description

	Dataset 1	**Dataset 2**	**Dataset 3**
Identified treatment	Promotion	Overtime	Coach
Treatment size (*T*)	319	1054	683
Control size (*C*)	14680	416	314
Treatment correlation with target variable (i.e. Employee Turnover)	6.18%	24.61%	4.84%

After splitting every dataset into training data and testing data, XGB is used to train the models with the data size of 10499, 1029, and 570 on dataset 1, 2, and 3, respectively. Predictive evaluation: Table 2 presents the prediction result — i.e. predicting the testing data with the trained model — of the three datasets. Note that the confusion matrices and accuracy results from positive and negative treatment correlations are combined to make it easier to observe the result and yielding a two-fold confusion matrix as shown in Table 2. After ranking the employees by the prescription result, the uplift value captured by the ETP and ETU models in every segmentation is calculated with LGWUM as visualized in Figure 4. Furthermore, the Qini curves visualize the ideal proportion of employees to be treated with the employee retention program. Figure 4 shows the target distributions captured by the ETP model and ETU model in the top 30% deciles. Lastly, the Jupyter Notebook file of this research is shared as the supplementary material of this research paper.

Table 2. Two-fold confusion matrix and accuracy result

			Dataset 1 ($N = 2 \times 4500$)		**Dataset 2 ($N = 2 \times 441$)**		**Dataset 3 ($N = 2 \times 245$)**	
			True	**False**	**True**	**False**	**True**	**False**
ETP model		**Positive**	6799	73	742	14	149	99
		Negative	121	2007	105	21	87	155
		Average Accuracy	97.84%		86.39%		65.84%	
ETU model	***CN***	**Positive**	6815	71	808	16	437	7
		Negative	122	1992	55	3	37	9
	CR	**Positive**	2056	250	666	24	449	7
		Negative	71	6623	185	7	28	6
	TN	**Positive**	8986	0	812	2	190	104
		Negative	7	7	66	2	71	125
	TR	**Positive**	8815	19	152	300	178	98
		Negative	128	23	281	536	80	134
	Average Accuracy		96.37%		62.25%		51.48%	

VI. Discussion

A. Identified treatment

Two employee retention programs for each dataset are successfully applied to solve employee turnover as visualized in Figure 4. Experiment A1–A3 yielding positive curves — pointing upwards — due to the tendency of the treatment with a *negative correlation* to reduce the employee turnover. On the contrary, experiment B1–B3 yielding negative curves — pointing downwards — due to the tendency of the treatment with a *positive correlation* to increase the employee turnover. Furthermore, experiments A3 and B3 in Figure 4 shows that despite a treatment correlation of 4.84% ($p < 0.05$), the ETU model is still capable to give good results.

FINDINGS: Two kinds of employee retention program may be identified and successfully applied by using the treatment identification approach for the HR dataset with an unidentified treatment. Moreover, the identified treatment may be successfully applied even with a correlation to employee turnover of less than 5%.

B. Predictive performance

The predictive evaluation results of the three datasets show that by average the ETP model yields 83.35% accuracy while the ETU model only yields 70.03% accuracy. The ETP models are trained with the objective of predicting employee turnover while ETU models are trained with the objective of prescribing employees that should be targeted with a retention program. Therefore, it is reasonable that ETP models outperform ETU models' predictive performance. Note that ETU models are multi-label binary classification, thus the accuracy function from Scikit–learn computes *subset accuracy* for ETU models. *Subset accuracy* does not count the true negative as accuracy's percentage while for binary classifications the true negative is counted as accuracy's percentage.

We speculate that the predictive performances for the ETP model or ETU model might be improved by applying several techniques that are not applied in this research to improve predictive performance, for instance, hyper–tuning or feature engineering. Moreover, Dataset 1 has the biggest training data and has the highest accuracy, followed by Dataset 2, and finally Dataset 3 with the smallest training data and lowest accuracy. Therefore, we also speculate that bigger training data will improve predictive performance for both models.

FINDINGS: The employee turnover prediction model is consistently yielding better predictive results than the employee turnover uplift model.

C. Prescriptive performance and target class distribution

As explained in Section C of Methodology, a model is considered to successfully target the right employees if the qini coefficient is greater than 5% (P>0.05). Therefore, the Qini coefficient results in Table 3 clearly show that the ETP model only successfully targets the right employees in three experiments — i.e. 50% success rate —. Intuitively, using the ETP model as prescriptive analytics by targeting employees with a high risk of turnover is a sensible approach, but the empirical evidence in this research shows that this approach which has been widely researched before is not reliable in targeting employees with a treatment. On the other hand, rather than assuming that the employees with a high turnover risk will be retained if targeted with a retention program, the ETU model specifically measures the ideal employees — i.e. Persuadables — to be targeted with a retention program. The ETU model introduced in this research is successfully targeting the right employees in all of the six experiments in this research — i.e. 100% success rate.

Table 3. Qini coefficient

		Dataset 1	**Dataset 2**	**Dataset 3**
Treatment correlation < 0	**ETP model**	-44.35%	35.91%	-12.92%
	ETU model	765.81%	23.99%	20.66%
Treatment correlation > 0	**ETP model**	302.03%	-41.58%	13.00%
	ETU model	720.47%	18.91%	14.17%

Moreover, Figure 4 visualizes the Qini curves that can be used to observe how the models prescribe the employees in each segmentation. As visualized in Figure 4, the ETU model performs better in experiments A1, A3, and B2, this indicates that the ETU model successfully separates between the Persuadables and Do-Not-Disturb employees. Targeting the persuadables will causing the Qini curve to point upwards while targeting the Do-not-Disturb will yield the opposite result. In experiment B1, the ETU model only yields good results around the first 50% deciles. This is as expected because in the real case scenario, the organization may only target a small number of employees with ERP to save expenses. Moreover, the uplift values and employee proportions are shown in percentage, for instance, a 0.2 uplift value in 0.25 employee proportion means the organization will get 20% of uplift by just targeting the top 25% of employees with the ERP.

As explained earlier in Section C of Literature, the preferable uplift class to be targeted by the retention program is the CNTR — i.e. *Persuadables*. Therefore, an optimum result may be achieved by only targeting CN or TR employees and avoiding CR or TN employees. Nevertheless, swarm plot D3 in Figure 5 shows that the ETP model incorrectly ranks the *TN* employees in the top 30% decile of the three datasets. On the other hand, swarm plots C1 and D4 clearly show that the ETU model correctly ranks the *CN* and *TR* employees, respectively. Despite the consistent performance of the ETU model, swarm plot C2, C3, D2, and D3 show that the ETU model still incorrectly ranks several CR and TN employees in the top deciles. Therefore, the Qini curve that projected by the ETU model tends to be decreased in top deciles, as shown in experiments A1, B2, and B3. These mistakes could be crucial if an organization has a limited budget for an employee retention program, and only capable to target a small proportion of employees. We speculate that these mistakes may be reduced by applying several techniques that are not applied in this research to improve prescriptive performance, for instance, hyper–tuning or other uplift modeling strategies.

Moreover, Experiments A1, A2, and A3 yielding a maximum uplift of ≈40%, ≈20%, and ≈10%, respectively. Therefore, the ETU model of Experiment A1 has both the highest maximum uplift and accuracy, followed by A2, and finally A3 with the lowest maximum uplift and accuracy. This indicates that higher accuracy may lead to higher maximum uplift for the ETU model with treatment that negatively correlated with employee turnover.

FINDINGS: Applying the Employee Turnover Prediction model as prescriptive analytics is not reliable, and could give a negative result. By contrast, applying the Employee Turnover Uplift model as prescriptive analytics is more reliable.

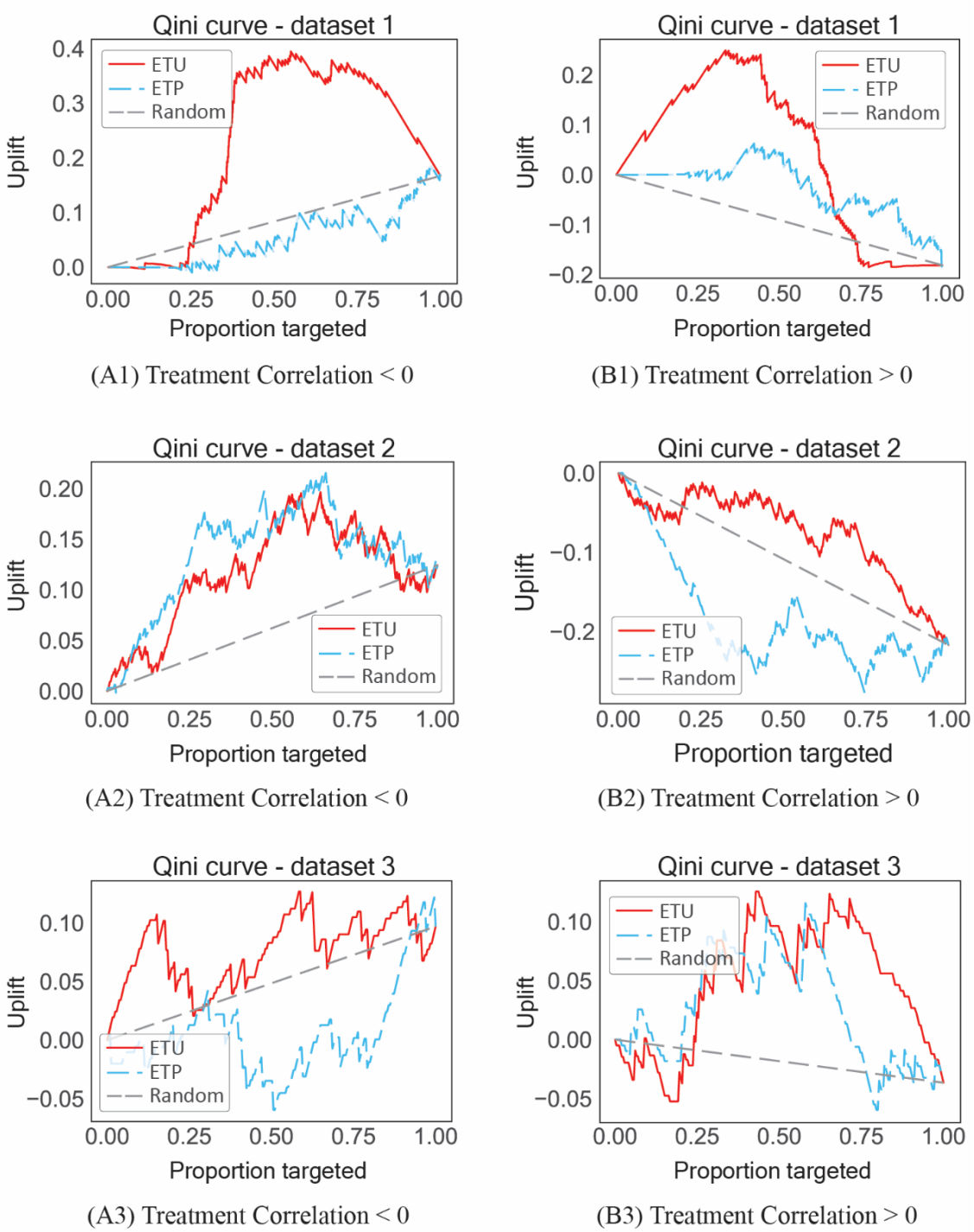

Figure 4. Prescriptive evaluation results with the Qini curve that represents the uplift value obtained by the model in the proportion of targeted employees. Red, blue, and gray curves, representing ETU, ETP, and random model, respectively. The results with negative treatment correlation are shown on the left, while the results with positive treatment correlation are shown on the right. (Please access the digital version of this manuscript to distinguish the color)

D. Future research

This research opens up four directions for future research about HR Prescriptive Analytics. *Firstly*, a comparative study between various uplift strategies still needs to be explored to find out the most reliable uplift strategy. *Secondly*, various techniques to improve performance also need to be explored,

 ISSN: 2579-7298

for instance, various machine–learning algorithms, hyper–tuning, feature engineering, etc. *Thirdly*, other techniques to measure the prescriptive performance may be explored, for instance, by comparing the Qini coefficient or measuring the uplift value in the top deciles.

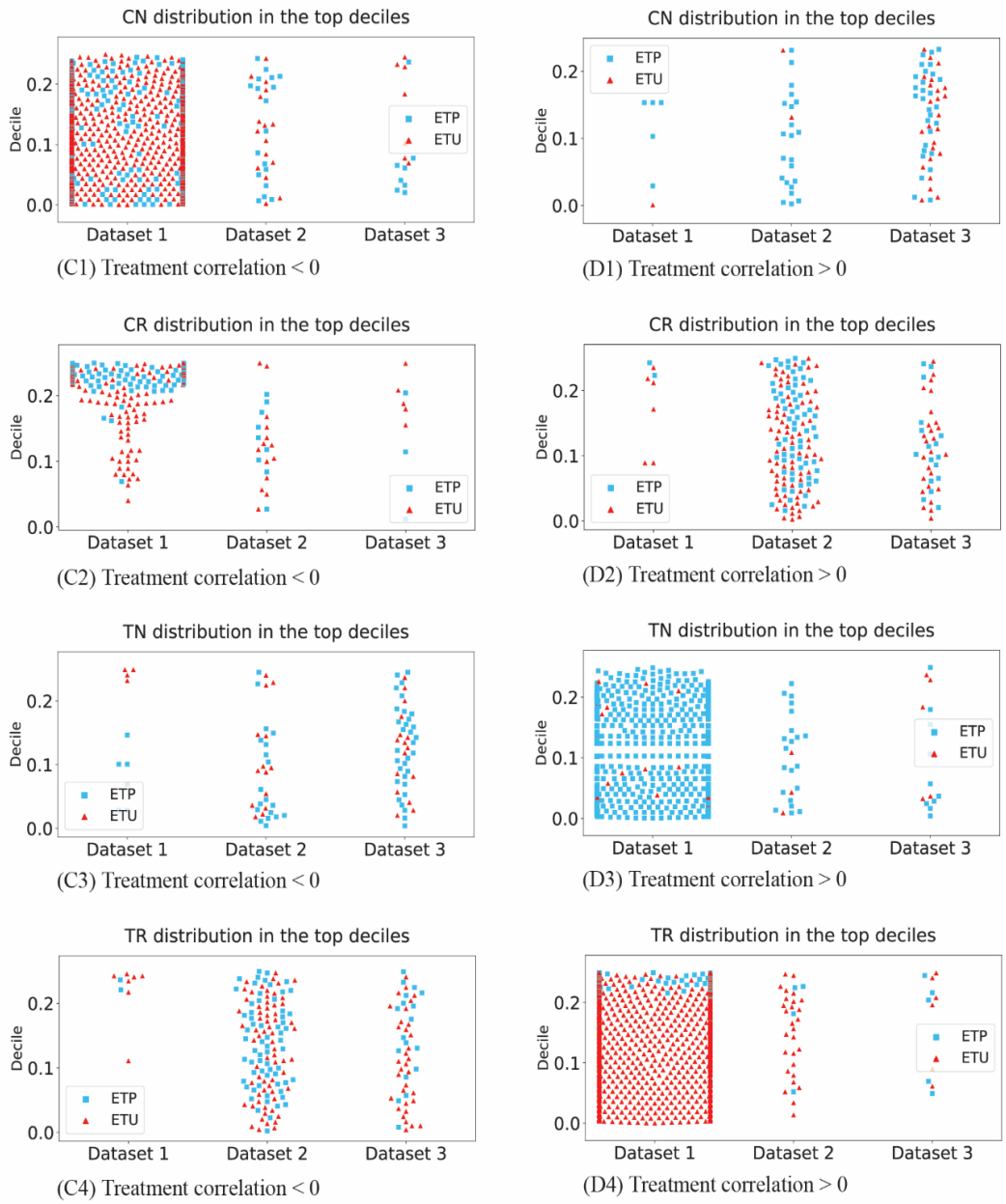

Figure 5. The four target classes recorded by the models are evaluated with swarm plots on the ranked employee data. The blue square represents employee data from the ETP model, while the red triangle represents employee data from the ETU model. The results with negative treatment correlation are shown on the left, while the results with positive treatment correlation are shown on the right. (Please access the digital version of this manuscript to distinguish the color)

Fourthly, despite the uplift value measured in this research, a profit-centric approach needs to be explored in measuring the prescriptive performance of an employee retention program. This approach may be conducted by applying the cost and benefit analysis that has been widely studied for the customer retention program. *Last but not least*, more empirical analysis is needed to be studied with other HR data or in a broader variety of applications — e.g. employee survival analysis, medicine recommendation, multi-treatment uplift, etc.

VII. Conclusion

We have developed a treatment identification technique to unlock the prescriptive capability of human resources data with unidentified treatment. This approach results in two possible prescriptive scenarios for each treatment identified, the first one has a negative correlation with employee turnover, and the other one has a positive correlation with employee turnover. Moreover, we successfully apply the conventional predictive model and the uplift model to three datasets with unidentified treatment to test the merits of this approach.

Empirical evidence of the uplift model outperforming the conventional predictive model is growing steadily in the literature. However, similar studies have never been conducted on human resources data. Therefore, this research presents an experimental case study with synthetic and real data to assess the performance comparison of the uplift model and conventional predictive model. While conducting the experimental case study, we also introduce a combination of *extreme gradient boosting* and *Lai's generalized weighed uplift* to predict and prescribe employee turnover. The result indicates that the uplift model is more reliable than the conventional predictive model in prescribing the persuadables employee that should be treated to reduce employee turnover. Therefore, this result also validates previous works about the merits of uplift modeling as prescriptive analytics.

Acknowledgment

We thank Giri Pujar, Pavan Subhash, and Eduard Babushkin for sharing the datasets that were studied in this research. We also thank Evy Romboult and Marie-Anne Guerry from Vrije Universiteit Brussel for sharing their early full-text publication, namely "The effectiveness of employee retention through an uplift modeling approach". Last but not least, we also thank the anonymous reviewers who critically reviewing this manuscript and providing meaningful suggestions.

References

[1] P. W. Hom, T. W. Lee, J. D. Shaw, and J. P. Hausknecht, "One hundred years of employee turnover theory and research," *J. Appl. Psychol.*, vol. 102, no. 3, pp. 530–545, 2017, doi: 10.1037/apl0000103.

[2] A. A. Kumar and K. B. Mathimaran, "Employee Retention Strategies â€"An Empirical Research," *Glob. J. Manag. Bus. Res.*, 2017.

[3] K. Simbeck, "HR analytics and ethics," *IBM J. Res. Dev.*, vol. 63, no. 4/5, pp. 1–9, 2019.

[4] K. Lepenioti, A. Bousdekis, D. Apostolou, and G. Mentzas, "Prescriptive analytics: Literature review and research challenges," *Int. J. Inf. Manage.*, vol. 50, pp. 57–70, 2020.

[5] L. Berk, D. Bertsimas, A. M. Weinstein, and J. Yan, "Prescriptive analytics for human resource planning in the professional services industry," *Eur. J. Oper. Res.*, vol. 272, no. 2, pp. 636–641, 2019.

[6] J. Kaur and K. S. Mann, "AI Based HealthCare Platform for Real Time, Predictive and Prescriptive Analytics," in *International Conference on Computing, Analytics and Networks*, 2017, pp. 138–149.

[7] A. Diez-Olivan, J. Del Ser, D. Galar, and B. Sierra, "Data fusion and machine learning for industrial prognosis: Trends and perspectives towards Industry 4.0," *Inf. Fusion*, vol. 50, pp. 92–111, 2019.

[8] D. R. S. Kamath, D. S. S. Jamsandekar, and D. P. G. Naik, "Machine Learning Approach for Employee Attrition Analysis," *Int. J. Trend Sci. Res. Dev.*, vol. Special Is, no. Special Issue-FIIIIPM2019, pp. 62–67, 2019, doi: 10.31142/ijtsrd23065.

[9] S. Yadav, A. Jain, and D. Singh, "Early Prediction of Employee Attrition using Data Mining Techniques," *Proc. 8th Int. Adv. Comput. Conf. IACC 2018*, pp. 349–354, 2018, doi: 10.1109/IADCC.2018.8692137.

[10] S. Saranya and J. S. Devi, "Predicting Employee Attrition Using Machine Learning Algorithms and Analyzing Reasons for Attrition," *Int. J. Adv. Eng. Res. Technol.*, vol. 6, no. 9, pp. 475–478, 2018, [Online]. Available: www.ijaert.org.

[11] I. W. Ariawan, "Predictive analysis of employee turnorver: a comparative study using logistic regression and artificial neural network," *J. Indones. Math. Soc.*, vol. 25, no. 3, pp. 325–335, 2019.

[12] K. G. King, "Data analytics in human resources: A case study and critical review," *Hum. Resour. Dev. Rev.*, vol. 15, no. 4, pp. 487–495, 2016.

[13] K. Sehgal, H. Bindra, A. Batra, and R. Jain, "Prediction of Employee Attrition Using GWO and PSO

Optimised Models of C5. 0 Used with Association Rules and Analysis of Optimisers," in *Innovations in Computer Science and Engineering*, Springer, 2019, pp. 1–8.

[14] S. S. Alduayj and K. Rajpoot, "Predicting Employee Attrition using Machine Learning," in *2018 International Conference on Innovations in Information Technology (IIT)*, 2018, pp. 93–98.

[15] Y. Zhao, M. K. Hryniewicki, F. Cheng, B. Fu, and X. Zhu, "Employee turnover prediction with machine learning: A reliable approach," in *Proceedings of SAI intelligent systems conference*, 2018, pp. 737–758.

[16] J. Frierson and D. Si, "Who's Next: Evaluating Attrition with Machine Learning Algorithms and Survival Analysis," in *International Conference on Big Data*, 2018, pp. 251–259.

[17] D. Bertsimas and N. Kallus, "From predictive to prescriptive analytics," *Manage. Sci.*, vol. 66, no. 3, pp. 1025–1044, 2020, doi: 10.1287/mnsc.2018.3253.

[18] M. Ramannavar and N. S. Sidnal, "A proposed contextual model for big data analysis using advanced analytics," *Adv. Intell. Syst. Comput.*, vol. 654, pp. 329–339, 2018, doi: 10.1007/978-981-10-6620-7_32.

[19] L. Berk, D. Bertsimas, A. M. Weinstein, and J. Yan, "Prescriptive analytics for human resource planning in the professional services industry," *Eur. J. Oper. Res.*, vol. 272, no. 2, pp. 636–641, 2019, doi: 10.1016/j.ejor.2018.06.035.

[20] D. Pessach, G. Singer, D. Avrahami, H. C. Ben-Gal, E. Shmueli, and I. Ben-Gal, "Employees recruitment: A prescriptive analytics approach via machine learning and mathematical programming," *Decis. Support Syst.*, p. 113290, 2020.

[21] E. Rombaut and M.-A. Guerry, "The effectiveness of employee retention through an uplift modeling approach," *Int. J. Manpow.*, Dec. 2019.

[22] D. Olaya, K. Coussement, and W. Verbeke, "A survey and benchmarking study of multitreatment uplift modeling," *Data Min. Knowl. Discov.*, vol. 34, no. 2, pp. 273–308, 2020, doi: 10.1007/s10618-019-00670-y.

[23] F. Devriendt, D. Moldovan, and W. Verbeke, "A Literature Survey and Experimental Evaluation of the State-of-the-Art in Uplift Modeling: A Stepping Stone Toward the Development of Prescriptive Analytics," *Big Data*, vol. 6, no. 1, pp. 13–41, 2018, doi: 10.1089/big.2017.0104.

[24] F. Devriendt, J. Berrevoets, and W. Verbeke, "Why you should stop predicting customer churn and start using uplift models," *Inf. Sci. (Ny).*, 2020, doi: 10.1016/j.ins.2019.12.075.

[25] E. Ascarza, "Retention futility: Targeting high-risk customers might be ineffective," *J. Mark. Res.*, vol. 55, no. 1, pp. 80–98, 2018, doi: 10.1509/jmr.16.0163.

[26] S. N. Mishra, D. R. Lama, and Y. Pal, "Human Resource Predictive Analytics (HRPA) For HR Management In Organizations," *Int. J. Sci. Technol. Res.*, vol. 5, no. 5, pp. 33–35, 2016.

[27] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 2016, pp. 785–794.

[28] D. Nielsen, "Tree boosting with xgboost-why does xgboost win" every" machine learning competition?" NTNU, 2016.

[29] N. J. Radcliffe, "Using control groups to target on predicted lift: Building and assessing uplift model," *Direct Mark. Anal. J.*, no. 3, pp. 14–21, 2007.

[30] R. Gubela, A. Bequé, S. Lessmann, and F. Gebert, "Conversion uplift in e-commerce: A systematic benchmark of modeling strategies," *Int. J. Inf. Technol. Decis. Mak.*, vol. 18, no. 03, pp. 747–791, 2019.

[31] K. Kane, V. S. Y. Lo, and J. Zheng, "Mining for the truly responsive customers and prospects using true-lift modeling: Comparison of new and existing methods," *J. Mark. Anal.*, vol. 2, no. 4, pp. 218–238, 2014.

[32] D. J. Benjamin *et al.*, "Redefine statistical significance," *Nat. Hum. Behav.*, vol. 2, no. 1, p. 6, 2018.