

СЕКЦІЯ 1

МОДЕЛЮВАННЯ, АНАЛІЗ ТА ОПТИМІЗАЦІЯ СКЛАДНИХ СИСТЕМ

УДК 004.9

Беляєв О.Р.¹, Наук.керів. – Коряшкіна Л.С.²

КОРОТКОСТРОКОВЕ ПЕРЕДБАЧЕННЯ ЗВІЛЬНЕНЬ ПРАЦІВНИКІВ МАСОВИХ ПРОФЕСІЙ В РІТЕЙЛ-КОМПАНІЇ

Високі показники зміни кадрів можуть зашкодити продуктивності організації та потенційно порушити інвестиції в людські ресурси. Зазвичай організації не готові до передчасної відставки працівників, і проблеми, що виникають внаслідок зміни кадрів, у цьому випадку ще більш шкідливі. Оцінка схильності співробітників до звільнення є першорядним для зменшення зміни кадрів, зменшуючи тим самим його негативний вплив на результати діяльності організації.

В даній роботі вирішується питання по короткостроковому передбаченню добровільного звільнення працівників масових професій в ритейл-компанії за допомогою регресії Кокса. В процесі досліджень виявлені і проаналізовані фактори, котрі зменшують або збільшують ризик звільнення. У роботі застосовувались бібліотеки `lifelines` та `pandas` мови програмування Python.

Для навчання моделі були зібрані дані по наразі вже звільненим продавцям та касирам, починаючи з 2018 року. Вибірка включала також фактори, які тим чи іншим чином можуть впливати на кількість днів роботи на посаді. Для швидкого доступу до інформації та її подальшого оновлення була створена вітрина даних.

Через шум в даних, нерівне співвідношення кількості записів при різних рівнях якісних факторів та велику плінність кадрів на ранніх етапах роботи співробітника, були вилучені всі записи, які негативно впливають на якість прогнозу моделі. Після ретельного аналізу у вибірці залишилися лише ті записи, в яких співробітники працювали більше 30 днів та не мали нульових значень певних предикторів.

На початку роботи з даними було виявлено і вилучено колінеарні фактори. Розрахований коефіцієнт кореляції Пірсона між кожними парами кількісних факторів. У випадку, коли значення коефіцієнту кореляції між двома будь-якими предикторами більше за 0.8 або нижче за -0.8, одну з двох змінних було вилучено з вибірки.

Модель будувалася з метою максимізації функції логарифмічного вподобання. За допомогою перехресної валідації (`cross-validation`) була обрана краща модель за рахунок підбору двох вхідних параметрів: штрафу до розміру коефіцієнтів під час регресії та співвідношенню для призначення штрафу L1 проти L2. Найкращі результати досягнуто при нульових значеннях `l1_ratio` та

¹ Беляєв О.Р. – студент НТУ «Дніпровська політехніка»

² Коряшкіна Л.С. – к.ф.м.н., доцент кафедри системного аналізу і управління НТУ «Дніпровська політехніка»

penalizer. Модель регресії за таких значень параметрів має індекс конкордантності, що дорівнює 0.751.

За результатами досліджень, факторами, котрі збільшують ризик звільнення, виявились: процентна кількість штрафів за крайній місяць роботи; процент відредагованого робочого часу у меншу сторону; кількість оформлених відпусток за крайні 3 місяці роботи; кількість оформлених лікарняних за крайній місяць роботи та зменшення ставки у меншу сторону за крайні 3 місяці роботи. А факторами, котрі знижують ризик, є: кількість оформлених оплачуваних відпусток на крайній посаді; збільшення ставки у більшу сторону за крайні 3 місяці роботи та кількість переміщень всередині компанії (кількість змін посади).

Для передбачення співробітників, котрі ймовірніше всього звільняться у найближчі 30 днів, використана функція виживання. Оцінка підраховувалась наступним чином: на вхід до моделі подається список всіх співробітників масових професій, котрі наразі працюють, та їх предиктори, на основі яких здійснюється прогноз. Далі відсікаються дні з прогнозованим значенням функції виживання, період котрих перевищує зазначені 30 днів. По кожному співробітнику беремо мінімальне значення, та відсікаємо 95% перцентиль серед усіх значень по всім записам, котрі подані на вхід. Ті співробітники, значення функції виживання у котрих більше за значення перцентіля, є претендентами на ймовірне добровільне звільнення у найближчі 30 днів. Отримані результати прогнозу можуть бути корисними для співробітників департаменту HR, які розробляють або заохочувальні заходи для подальшого затримання цінного співробітника, або заздалегідь підбирають нового робітника на відповідну посаду.

Розроблений програмний продукт швидко перенавчає модель на нових даних і може бути адаптований під потреби бізнесу, використовуючи нові підходи і алгоритми виживання.

1. Сайт Lifelines [електронний ресурс]: Introduction to survival analysis. URL: <https://lifelines.readthedocs.io/en/latest/Survival Analysis intro.html>