

УДК 51-77 : 331.108.242

Н.О. Манакова¹, А.А. Євдокімов², В.П. Манаков³, І.І. Циганенко⁴¹ ХНУМГ ім. О. М. Бекетова, м. Харків, Україна, natalymanakova@gmail.com² ХНУМГ ім. О. М. Бекетова, м. Харків, Україна, evdokimo-andre@yandex.ru³ ХНУРЕ, м. Харків, Україна, volodymyr.manakov@knure.ua⁴ ХНУМГ ім. О. М. Бекетова, м. Харків, Україна, macnat@rambler.ru

ДИЗАЙН ФРЕЙМВОРКУ АНАЛІЗУ ДАНИХ ЩОДО ПЛИННОСТІ КАДРІВ МЕТОДАМИ DATA MINING

В роботі проведено огляд методів Data Mining, для рішення кадрових питань, зокрема прогнозування «наміру звільнитися». Розуміння того які з факторів вплинули на рішення працівника звільнитися, або навпаки залишитися, незважаючи на те, що він мав намір звільнитися, дуже важливо для вибудовування кадрової політики підприємства. В цій роботі ми зосереджені на прогнозуванні «наміру звільнитися», що є скоріше когнітивним чим кількісним поняттям, тим самим ускладнюючи математичний вимір. Розроблений дизайн фреймворка дослідження готовності до звільнення за власним бажанням економічно активних верств населення. Цей фреймворк призначений для вивчення взаємозв'язку між соціальним обміном і особистістю в прогнозуванні наміру звільнитися.

НАМІР ЗВІЛЬНИТИСЯ, ОБОРОТ КАДРІВ, ПЛИННІСТЬ КАДРІВ, РУХ ПЕРСОНАЛУ, DATA MINING

Манакова Н.О., Євдокімов А.А., Манаков В.П., Цыганенко И.И. Дизайн фреймворка анализа данных о текучести кадров методами Data Mining. В работе проведен обзор методов Data Mining для решения кадровых вопросов, в частности прогнозирования «намерения уволиться». Понимание того какие из факторов повлияли на решение работника уволиться, или наоборот остаться, несмотря на то, что он хотел уволиться, очень важно для построения кадровой политики предприятия. В этой работе мы сосредоточены на прогнозировании «намерения уволиться», что является скорее когнитивным чем количественным понятием, тем самым усложняя математические измерения. Разработан дизайн фреймворка исследования готовности к увольнению по собственному желанию экономически активных слоев населения. Этот фреймворк предназначен для изучения взаимосвязи между социальным обменом и личностью в прогнозировании намерения уволиться.

НАМЕРЕНИЕ УВОЛИТЬСЯ, ОБОРОТ КАДРОВ, ТЕКУЧЕСТЬ КАДРОВ, ДВИЖЕНИЕ ПЕРСОНАЛА, DATA MINING

Manakova N.O., Evdokimov A.A., Manakov V.P., Tsyganenko I.I. The framework for analyzing data on the turnover of personnel using Data Mining. In the work, the review of Data Mining methods for personnel issues, in particular, the forecasting of the “intention to quit”, was conducted. The design of the framework for the study of turnover readiness of economically active parts of the population has been developed. This framework is consigned to study the relationship between social issued and personality features under staff turnover prediction.

TASK IS OPENED, BUSINESS TRAFFICKING, LOAD CARRIERS, PERSONAL PERSONNEL, DATA MINING

Вступ

Висока цінність деяких співробітників для успішного функціонування підприємства призводить до необхідності зниження ризику звільнень за власним бажанням, що може призводити до високої плинності персоналу. Негативні переміщення коштують організації значних фінансових витрат. Це втрачені можливості, втрата клієнтів, цінних працівників (не рівноцінні заміни), витрати на заміну, підбір, адаптацію персоналу, навчання нових працівників, втрати фонду оплати праці пов'язані з рухом персоналу, втрати пов'язані зі зменшенням продуктивності та віддачі від персоналу, відповідно зниження продуктивності і роботи команд [1]. Соціальні наслідки зниження рівня привабливості організації як роботодавця та довіри до неї, як зі сторони потенційних кандидатів, так і споживачів/клієнтів та морального рівня працівників, команд та всього колективу.

Кадрові рішення, пов'язані з рухом персоналу (найчастіше спрямовані запобіганню плинності) приймаються в більшості випадків суб'єктивно, без аналізу статистичних даних, їх динаміки, тенденцій та закономірностей руху персоналу. Однією з причин такого стану є відсутність теоретичних та практичних напрацювань які описують систему та процеси руху персоналу. А також інструментів його аналізу, зокрема повної системи показників оцінки руху персоналу. Мобільність персоналу, його плинність, тимчасове та постійне вивільнення персоналу представляє собою доволі різнобічне поняття, що розглядається як багатофакторний процес. Враховуючи дані тенденції розвитку науки управління персоналом та підходи до розуміння її системи, функцій, технологій, механізмів реалізації та інструментарію.

1. Актуальність та постановка задач

Одним з найбільш поширених методів аналізу плінності кадрів, є застосування цілого ряду різноманітних коефіцієнтів або показників. В контексті дослідження управління рухом персоналу організації, зростає актуальність питання аналізу існуючих показників руху персоналу організації, їх переосмислення та доповнення.

Аналітичне опрацювання згаданих нижче показників дає можливість бачити їх динаміку, відносно конкретної організації (ефект від процесів руху персоналу та управлінських рішень пов'язаних з ними), мати можливість періодично порівнювати їх з загальними та локальними даними по ринку, відслідковувати тенденції в сегменті в якому працює організація та загальноринкові. Це створює можливість розуміння місцезнаходження організації по відношенню до інших як в своєму сегменті, та і в загальноринковому. Окрім того, це дає можливість знаходити свою траєкторію руху та розвитку, та є інструментом планування роботи з персоналом.

Але, аналіз плінності кадрів за допомогою показників дає знеособлене уявлення, і не враховує не тільки особисті характеристики працівника, а й причини, через які він прийняв рішення піти з роботи. В цій роботі ми будемо зосереджені на прогнозуванні «наміру звільнитись», що є скоріше когнітивним чим кількісним поняттям, тим самим ускладнюючи математичний вимір. Однак багатьма дослідниками були розроблені різноманітні конструкції для вимірювання цього наміру з багатообіцяючими результатами.

Наприклад, 2000 року Роджер Гріффет і Пітер Хом [2] провели метааналітичні дослідження (дослідження проведених в цій галузі досліджень), в ході якого з'ясувалося «намір звільнитися» реалізується фактичним звільненням приблизно на 12 % у всіх дослідженнях. .

Проте розуміння того які з факторів вплинули на рішення працівника звільнитися, або навпаки залишитися, незважаючи на те, що він мав намір звільнитися, дуже важливо для вибудовування кадрової політики підприємства.

Для вивчення згаданих чинників розроблений дизайн фреймворка дослідження готовності до звільнення за власним бажанням економічно активних верств населення. Цей фреймворк призначений для вивчення взаємозв'язку між соціальним обміном і особистістю в прогнозуванні наміру звільнитися персоналу. Необхідність проведення якісного дослідження полягає в тому, щоб виявити найбільш важливі фактори, які будуть використані для прогнозування наміру звільнитися, після чого кількісні методи будуть використані для вивчення ступеня впливу найбільш важливих факторів і розробки моделі прогнозування.

У наступному розділі розглядаються різні аспекти дизайну описаного фреймворка.

2. Дизайн опитування

Цільова група дослідження: економічно активне населення України віком 15–70 років. Надійність оцінювання показників за результатами обстеження, планується забезпечити дизайном репрезентативної вибірки побудованої на основі імовірнісного багатоступеневого стратифікованого кластерного відбору населених пунктів і домогосподарств, а також випадкового відбору респондентів у межах відібраних домогосподарств.

Економічно активне населення включає частину населення обох статей віком від 15 до 70 років включно, яке протягом певного періоду забезпечує пропозицію робочої сили на ринку праці.

За даними вибірових обстежень населення (домогосподарств) з питань економічної діяльності, все населення країни у віці 15-70 років розподіляється як три взаємовиключні та вичерпні категорії: зайняті, безробітні та економічно неактивні (які не беруть участі в ринку праці).

Економічно активне населення (робоча сила) включає тільки тих осіб, які займалися економічною діяльністю або шукали роботу і були готові почати працювати, т. Е. Класифіковані як зайняті або безробітні.

Рівень економічної активності є відносною мірою щодо частки робочої сили. Він визначається як відношення (у відсотках) до чисельності економічно активного населення у віці 15–70 років до загальної чисельності населення зазначеного віку чи населення відповідної соціальної та демографічної групи.

Зайнятість означає не тільки роботу за заробітну плату (найману працю), але також заради отримання прибутку або сімейного доходу на власному підприємстві, включаючи виробництво у власному підсобному сільському господарстві для продажу виробленої продукції на ринку. А також, якщо вони тимчасово були відсутні на роботі, тобто формально були прикріплені до місця роботи, мали своє власне підприємство (бізнес), але з певних причин не працювали протягом звітного періоду.

Зайняте населення не включає осіб, які виконували неоплачувану громадянську добровільну роботу, та осіб, які виконували тільки домашні обов'язки. Громадянська добровільна діяльність не є економічним видом діяльності і має три відмінні риси: А) вона повинна бути оплачена; Б) здійснюється з власної волі, за відсутності тиску, наприклад, при активній військовій службі; В) виконується для організації, спільноти або особи, що не має зв'язку з родичами, за межами власного домашнього господарства.

Всі показники, які будуть оцінюватися при проведенні опитування, розбиті на основні блоки, які в свою чергу розбиті на взаємопов'язані групи:

- Персональні характеристики (демографічні показники, професійні характеристики, фінансовий стан домогосподарств, досвід роботи, досвід

безробіття, задоволеність поточним робочим місцем).

- Характеристика підприємства (базові показники та соціальна політика підприємства),
- А також вихідна змінна, яка характеризує намагання працівника звільнитися.

Нижче представлені показники і категорії відповідей на них по кожній групі.

Блок персональних характеристик.

Перша група показників представляє собою демографічні показники, що включають тільки загальні характеристики, такі як вік, стать, родинний статус, тип поселення та наявність дітей.

- Вік (15-70 років)
- Стать
 - жінки;
 - чоловіки.
- Тип поселення
 - обласний центр;
 - інше місто;
 - село.
- Родинний стан
 - в шлюбі;
 - в шлюбі не був;
 - розлучений;
 - вдова/вдівець.
- Кількість дітей
 - досадковий вік (0–2 роки) ;
 - садовий вік (3–6 років) ;
 - вік початкової школи (7–10 років) ;
 - вік середньої школи (11–15) ;
 - випускники/абітурієнти (16–18).

Професійні характеристики включають найбільш критичні для отримання робочого місця, довгої та успішної праці на ньому з релевантними обов'язками та зарплатнею.

- Освіта
 - вища;
 - професійно-технічна;
 - повна середня;
 - початкова та базова середня;
- Професії
 - законодавці, вищі державні службовці, керівники, менеджери (управителі);
 - професіонали;
 - фахівці;
 - технічні службовці;
 - працівники сфери торгівлі та послуг;
 - кваліфіковані робітники сільського та лісового господарств, риборозведення та рибальства;
 - кваліфіковані робітники з інструментом;
 - робітники з обслуговування, експлуатації та контролю за роботою технологічного устаткування, складання устаткування та машин;
 - найпростіші професії.
- Досвід
 - відсутній (0);
 - початковий (1-3 років) ;

- впевнений (4–6) ;
- великий (7 і більше років).

Фінансовий стан домогосподарства критично важлива група показників для оцінювання мотивації співробітників залишатися на робочому місці та більш активно шукати нове місце роботи. Включимо до цієї групи самооцінку фінансового стану, ресурси забезпечення та способи покриття поточних витрат.

- Самооцінка фінансового стану домогосподарства
 - доволі заможне;
 - приблизно на середньому рівні по країні;
 - не бідне, але нижче від середнього рівня у країні;
 - бідне;
 - відмова від відповіді.
- Основні ресурси матеріального забезпечення
 - заробітна плата;
 - підприємницький дохід;
 - випадкові доходи;
 - стипендія;
 - допомога батьків, інших родичів;
 - заробітна плата чоловіка/дружини, співмешканця;
 - доходи від власності, банківських рахунків, цінних паперів;
 - соціальні виплати держави;
 - інше.
- Способами покриття поточних витрат (кілька варіантів відповідей)
 - витрачали заощадження;
 - продавали майно вашого домогосподарства;
 - брали кредит/и в офіційній фінансовій установі;
 - позичали кошти у друзів, знайомих чи неофіційних кредиторів;
 - отримували соціальні субсидії, соціальну грошову допомогу;
 - економили на витратах (наприклад, витрати на харчування, житло, охорону здоров'я, освіту) ;
 - працювали додатково, щоб більше заробити (підробляли) ;
 - інше.

Наступна група, що також має сильний вплив на психологічні причини залишатись на робочому місці це попередній досвід безробіття. До цієї групи включено причини відмови від активного пошуку роботи, тривалість пошуків роботи, види отриманої допомоги в центрі зайнятості.

- Причинами відмови від активного пошуку роботи
 - не знаю, як або де шукати роботу;
 - не можу знайти роботу відповідно до своїх навичок;
 - раніше шукав роботу, але не знайшов жодної;
 - надто молодий для того, щоб знайти роботу;
 - у даній місцевості / районі роботи немає.

- Тривалість пошуку роботи
 - менше ніж тиждень;
 - від 1 тижня до 1 місяця;
 - від 1 місяця до 3 місяців;
 - від 3 місяців до 6 місяців;
 - від 6 місяців до 9 місяців;
 - від 9 місяців до 1 року;
 - від 1 року до 2 років;
 - 2 роки та більше;
 - питання не ставилося.
- Види отриманої допомоги від центрів зайнятості
 - поради щодо того, як шукати роботу;
 - інформація про вакансії;
 - орієнтація стосовно можливостей щодо освіти та професійної підготовки;
 - включення до програм освіти або професійної підготовки;
 - включення до програми громадських робіт;
 - жоден.

Задоволеність робочим місцем є можливо найбільш важливим фактором, але доволі важким для вимірювання. В нашому дослідженні для вимірювання цих показників буде використовуватись шкала Лікерта (Likert), яка вимагає спеціальної обробки даних. Ця процедура буде розглянута нижче в розділі дизайну методів обробки даних.

- Відпрацьовані години за тиждень
 - не працював останній тиждень;
 - менше 10;
 - 10–19;
 - 20–29;
 - 30–39;
 - 40–49;
 - 50–59;
 - більше 60;
 - відсутня відповідь.
- Відпрацьовані місяці на цьому робочому місці (вказати кількість місяців)
- Рівень задоволення від роботи
 - дуже задоволені;
 - скоріше задоволені;
 - скоріше незадоволені;
 - дуже незадоволені.
- Відповідно та невідповідно кваліфікованих
 - надмірна кваліфікація;
 - недостатня кваліфікація;
 - відповідна кваліфікація.
- Рівень задоволення від зарплатні
 - дуже задоволені;
 - скоріше задоволені;
 - скоріше незадоволені;
 - дуже незадоволені.
- Рівень зарплатні відносно галузі та професії
 - вища;
 - достатня;
 - середня;
 - недостатня;
 - занизька.

Блок характеристик підприємства.

Наступний блок характеристик, це характеристики підприємств. Цей блок включає по-перше базові описові характеристики підприємства. Необхідно підкреслити важливість включення сектора (галузі) до дослідження, тому що плинність кадрів відрізняється для різних галузей. Так, наприклад, за даними статистичного управління України до трійки лідерів за рівнем плинності кадрів серед сфер економіки України в першому півріччі 2010 року увійшли рибне господарство та рибна ловля (25,1%), торгівля, ремонт автомобілів, побутових виробів та предметів індивідуального користування (22,7%), діяльність готелів і ресторанів (20,5%). Найменша «плинність» зафіксована в сфері охорони здоров'я і надання соціальної допомоги (5,4%), в освіті (5,9%) та державному управлінні (6,4%).

- Сектори (галузі)
 - сільське господарство, лісництво та рибальство;
 - промисловість;
 - будівництво;
 - оптова та роздрібна торгівля; ремонт автомобілів та мотоциклів;
 - транспорт, складське господарство, поштова та кур'єрська діяльність;
 - послуги комерційного сектору;
 - державні та соціальні послуги;
 - готелі та ресторани;
 - інформація та телекомунікації;
 - фінансова та страхова діяльність;
 - операції з нерухомим майном;
 - професійна, наукова та технічна діяльність;
 - діяльність у сфері адміністративного та допоміжного обслуговування;
 - державне управління й оборона; обов'язкове соціальне страхування;
 - освіта;
 - охорона здоров'я та надання соціальної допомоги;
 - мистецтво, спорт, розваги та відпочинок;
 - надання інших видів послуг;
 - діяльність домашніх господарств;
 - відсутня відповідь.
- Тип підприємства
 - бюджетна організація/установа;
 - державне підприємство;
 - недержавна/неприбуткова громадська організація;
 - приватне підприємство/фермерське господарство;
 - приватне домогосподарство;
 - посольство, міжнародна організація;
 - інше.
- Розмір підприємства
 - дуже маленьке (до 10 співробітників);
 - маленьке (до 100 співробітників);
 - середнє (до 500 співробітників);
 - велике (до 1000 співробітників);
 - дуже велике (більше 1000 співробітників).

Іншим компонентом блоку характеристик підприємств розглянемо соціальну політику підприємства

- Тип контракту
 - письмовий трудовий договір та/або запис у трудовій книжці;
 - письмовий цивільно-правовий / господарський договір про виконання робіт чи надання послуг;
 - усна домовленість (без офіційного оформлення).
- Тривалість контракту
 - необмежений термін
 - обмежений термін
- Причини строкового контракту
 - на період навчання на робочому місці, стажування;
 - випробувальний строк;
 - сезонна робота;
 - випадкова/поденна робота;
 - робота із заміщенням відсутнього працівника;
 - державна програма зайнятості;
 - особлива/конкретна послуга або завдання;
 - інша причина;
- Вид соціальних гарантій
 - транспорт або компенсація транспортних витрат;
 - харчування або компенсація витрат на харчування;
 - щорічна оплачувана відпустка;
 - оплачуваний лікарняний;
 - додаткове недержавне пенсійне страхування;
 - вихідна допомога / допомога при звільненні;
 - оплата понаднормової роботи;
 - медичне страхування;
 - премія / винагороди за високі результати;
 - підприємство здійснювало внески на обов'язкове державне соціальне страхування – єдиний соціальний внесок;
 - освітні або навчальні курси (освітні відпустки);
 - безпека праці / захисні засоби або захисний спецодяг;
 - заклади або послуги догляду за дітьми;
 - відпустка у зв'язку з вагітністю та пологами / батьківством.

Блок вихідної змінної

Вихідна змінна «Намір змінити роботу та його причини» винесена в окремий блок, тому що вона представляє собою предиктор що вивчається. Намір не обов'язково буде реалізований, як вже зазначалося вірогідність цього складає 12%, але наше завдання попередження звільнення, а не констатація факту звільнення. Нижче в таблиці наведені категорії відповіді.

- Намір змінити роботу та причини
 - не маю наміру змінити роботу;

- нинішня робота є тимчасовою;
- страх втратити нинішню роботу;
- щоб працювати більше часу з оплатою за нинішньою ставкою;
- щоб мати більшу оплату за годину;
- щоб працювати менше часу зі зменшенням оплати;
- щоб краще використовувати свою кваліфікацію / навички;
- щоб мати зручніший графік та витратити менше часу на поїздку;
- щоб покращити умови праці;
- перехід від задовільної тимчасової зайнятості;
- перехід від учнівства/стажування на робочому місці;
- перехід від задовільної самозайнятості;
- перехід від неоплачуваної роботи в сімейному бізнесі.

3. Дизайн методів обробки

Первинна обробка та описові статистики

На початковому етапі дослідження при зборі даних будуть застосовуватися методи описової статистики і візуалізації для виявлення первинних залежностей з урахуванням типу шкали вихідних даних. Найбільший інтерес представляє обробка даних задоволеності за шкалою Лікерта (часто помилково згадується як шкала Лайкерта), яка названа так на честь її розробника (Likert, 1932).

Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
(1)	(2)	(3)	(4)	(5)

Рис. 1. Шкала Лікерта

Первісна шкала Лікерта використовувала серію питань з п'ятьма варіантами відповіді: рішуче схвалити (1), схвалити (2), не визначиться (3), відхилити (4) і зовсім не схвалювати (5). Лікерт об'єднав відповіді із серії питань, щоб створити комбінований індекс із серії з чотирьох або більше питань Лікерт-типу (Likert - type items). Він не аналізував окремі питання, хоча такий підхід теж можливий і іноді використовується. При розробці дизайну опитування досліднику необхідно прийняти рішення скільки категорій відповідей вибрати. У той час як Лікерт використовував п'ятибальну шкалу, довжину шкали часто збільшують, щоб створити семибальною шкалою, додавши «дуже» до відповідних верхнім і нижнім кордонів п'ятибальною шкалами. Інший поширений випадок - це усічення категорій до парного числа категорій (зазвичай чотири), щоб виключити «нейтральний» варіант.

Щоб правильно аналізувати дані такого типу, потрібно розуміти шкалу вимірювань. Якщо варіантів відповіді, присвоїти номери, то буде отримано відношення «більше ніж». Однак, наскільки

більше це не мається на увазі. Внаслідок цього питання Лікерт-типу (Likert-type items) відносяться до порядкової шкалою (ordinal scales).

Як правило, середнє і стандартне відхилення є неприйнятними для порядкових шкал, так само як і будь-які параметричні аналізи, засновані на нормальному розподілі. Для аналізу цих даних підходять непараметричні процедури, засновані на ранзі, медіанної або діапазонної шкалою, а також вільні методи розподілу, такі як таблиці, частоти, таблиці спряженості і статистика χ^2 -квадрат. В якості методу дисперсійного аналізу, прийнятним для порядкової шкали вважається метод Крускал-Уолліса.

З іншого боку, дані шкали Ликерта (тобто комбіновані індекси за серією питань) аналізуються на шкалі інтервальних вимірювань (interval scales). Але для перевірки гомогенності і узгодженості необхідно піддавати такі комбінації тесту Капі або обчисленню альфи Кронбаха. У разі, якщо дані пройшли зазначені тести, і розглядаються як інтервальні, можуть бути застосовані описові статистики на основі середніх і стандартного відхилення, а також параметричні тести такі як критерій Пірсона, ANOVA, t-test і деякі регресивні методи.

Просунуті методи аналізу

Будемо називати ці методи просунутими, оскільки вони вимагають спільного використання статистичних базових підходів і більш складних (але більш результативних) методів з розділу Data Mining.

Когортний аналіз

Когортний аналіз - це метод, який найчастіше використовується в маркетинговій аналітиці для відстеження відтоку клієнтів. Вони хочуть знати, коли клієнти перестануть користуватися їх сервісом. Аналогічно фахівець з кадрів хоче мати можливість знати момент часу, коли співробітники збираються звільнитися. Сегментуємо співробітників в когорти, ґрунтуючись на тому, коли вони надійшли на підприємство комбінуючи з додатковими показниками, що надає можливість відстежувати коефіцієнт утримання за періодами для конкретної когорти. Провівши спільний аналіз когорт можна побачити життєвий цикл робочого періоду співробітника.

Тобто, ви можете побачити, скільки з них зберігається з місяця в місяць. Подібно коефіцієнту утримання, перевага цього заходу полягає в тому, що вона відсікає частину шуму від змішаної вибірки з різних штатних співробітників. Усреднюючи набір даних вниз по стовпчику, ми можемо почати робити деякі (грубі) припущення про швидкість, з якою будь-які когорти співробітників зберігаються після того, як вони увійшли в компанію. Характеристика «грубі» вказує на те, що цей аналіз є скоріш приблизно понятійним, ніж засновані на статистичному аналізі.

Головна ідея когортного аналізу полягає в тому, що життєвий цикл співробітника впливає

на ймовірність його звільнення. Дослідження підтвердили, що нові співробітники (не обов'язково молоді) звільняються з більшою ймовірністю, ніж працівники з більш довгим часом найму (Nom et al 2008). Часто ця ж передумова ігнорується при використанні агрегованих базових показників, такі як коефіцієнт плинності або коефіцієнт утримання, а також виробити прогностичну міру для всього підприємства.

Когортний аналіз дозволяє нам почати аналізувати ймовірність звільнення, виходячи з того, де кожен співробітник знаходиться в життєвому циклі. Головним недоліком методу є те, що якщо ваша вибірка зроблена для занадто короткого часового періоду, може виявитися важким створити набір достатньої історичної глибини, щоб ця робота була ефективною.

Важливою перевагою методу є те, що хоча ви не можете отримати прогнози звільнення на індивідуальному рівні за допомогою цього аналізу, але придатний для планування робочої сили на агрегованому рівні. Хорошим прикладом когортного аналізу може стати більш точна оцінка того, скільки нових співробітників залишать роботу до кінця облікового періоду.

Логістична регресія

Логістична регресія - це метод, який часто зустрічається при створенні прогнозів обороту в різних сферах. Цей метод є одним із способів, за допомогою якого ми можемо почати розуміти такі питання, як «чому люди йдуть», «що ми можемо зробити, щоб вплинути на плинність кадрів» або «хто, швидше за все, піде в наступному році?».

Мета будь-якого типу регресії - передбачити результат з використанням одного або декількох інших факторів. В кадровому аналізі плинності ми повинні прогнозувати оборот кадрів як залежну змінну, використовуючи наші дані про робочі місця, які є незалежними змінними.

Особливістю логістичної регресії в тому, що вона є формою регресії, яка використовується, коли результат, який передбачається дорівнює 1 або 0. Для прогнозування плинності кадрового складу це дуже підходить, оскільки співробітник або звільняється, або ні.

Основним недоліком застосування логістичної регресії в сфері оцінки ризику втрати персоналу полягає в тому, що рівень точності прогнозу для людської поведінки не настільки високий як для технічних систем, де інтервал відмови машин набагато точніший і передбачуваний.

Ще одним недоліком логістичної регресії є досить нетривіальна інтерпретація, яка потребує кількох кроків. Спочатку визначається що одна група звільниться з більшою ймовірністю ніж інша, потім буде потрібно ще кілька кроків, щоб перетворити цей результат в абсолютну оцінку ризику

Аналіз виживаності

Аналіз виживаності, також відомий як аналіз історії подій, є вдосконаленим статистичним

методом, який використовується для оцінки ймовірності події, що відбувається з плином часу. Цей метод широко застосовується в медичних науках, де він використовувався для прогнозування виживаності пацієнтів, але може бути застосований і для оцінки плинності кадрів.

Так наприклад дослідження проведене Томом Бріггсом для американської асоціації HR, продемонструвало, додаток аналізу виживання для аналізу різниці у відсотку звільнення для кандидатів, яким давали реальну можливість попереднього ознайомлення з роботою і тих, яким не давали. Для першої групи кандидатів це призвело до зниження процента звільнень на 15%.

На відміну від вище розглянутих результатів логістичної регресії аналіз виживаності може бути використаний для отримання ймовірність звільнення в певний момент часу для конкретного співробітника.

Методи дерев — Древа прийняття рішень та Випадковий ліс (Decision Trees and Random Forest)

У порівнянні з логістичною регресією або аналізом виживання, які відразу приводять до індивідуальному рівню передбачення, модель дерева рішень починає з аналізу всіх співробітників, а потім сортує їх на менші і менші групи в залежності від ймовірності їх звільнення. Це створює деревоподібну структуру з центральним вузлом і багато листя для кожного можливого шляху від кореня (центрального вузла).

У своєму дослідженні [3] «Аналіз звільнення співробітників» Divyabh Misra він аналізує, як управлінська політика і інші фактори прийняті в компанії впливають на оборот в SanDisk.

Метод Random Forest ґрунтується на моделі дерева рішень наступним чином. На високому рівні випадковий ліс бере випадковий вибір даних з вашого набору даних і групує їх за своїми власними деревах рішень. Потім він бере середнє значення всіх створених дерев, щоб зробити прогноз. Ідея тут полягає в тому, що багато дрібніших передбачення, взяті разом, можуть в кінцевому підсумку створити сильніший прогноз.

Висновок

Інтелектуальний аналіз даних (Data Mining) є однією з нових технологій в області засобів аналізу даних, і має значний вплив в області людських ресурсів. Відповідний фреймворк по початковому збору даних дозволить отримати якісний набір для різних методів Data Mining. Це допоможе в розкритті закономірностей, щоб отримати знання, які допоможуть особам, які приймають рішення по розробці стратегій та планів на майбутнє ринку праці. Реалізація методів інтелектуального аналізу за даними опитувань та відділу кадрів можуть обробляти великі обсяги даних і перетворити його в корисну інформацію для поліпшення процесу

прийняття рішень в організаціях. Крім того, це може привести до більш легкому розумінню потреб організації та сформуванню працездатний підхід до прогнозування потреб ринку праці.

Список літератури: 1. Carifio, J. & Perla, R. (2007). Ten Common Misunderstandings, Misconceptions, Persistent Myths and Urban Legends about Likert Scales and Likert Response Formats and their Antidotes. *Journal of Social Sciences*, 2, 106-116. <http://www.scipub.org/fulltext/jss/jss33106-116.pdf> 2. Rodger W. Griffeth, Peter W. Hom, Stefan Gaertner. "Meta-Analysis of Antecedents and Correlates of Employee Turnover: Update, Moderator Tests, and Research Implications for the Next Millennium". *Journal of Management* 2000, Vol. 26, No. 3, 463-488 3. Divyabh Mishra "Employee Attrition Analysis Identify / Model key indicators of employee attrition for SanDisk employees based on publicly available data". Available online at 4. <https://www.slideshare.net/divyabh/employee-attrition-analysis>. checked on 4/1/2017 Richard Rosenow: Analyzing Employee Turnover. Descriptive Methods. Available online at <https://www.linkedin.com/pulse/analyzing-turnover-descriptive-methods-richard-rozenow-mp>, checked on 4/1/2017

Resume

N.O. Manakova, A.A. Evdokimov,
V.P. Manakov, I.I. Tsyganenko

THE FRAMEWORK FOR ANALYZING DATA ON THE TURNOVER OF PERSONNEL USING DATA MINING

Background: Personnel decisions related to the staff movement (most often aimed at preventing staff turnover) are taken in most cases subjectively, without analysis of statistical data, their dynamics, trends, and patterns of the staff movement. One of the reasons for this situation is the lack of theoretical and practical developments describing the system and processes of personnel movement. As well as tools for its analysis, in particular, a complete system of indicators for assessing the staff movement.

Materials and methods: One of the most common methods of analyzing staff turnover is the use of a number of different coefficients or indicators. In the context of the study of the organization's personnel management, the urgency of the issue of analyzing the existing turnover indices of the organization's personnel, their rethinking and supplementation are growing. To study the factors, the design of the framework for the study of turnover readiness of economically active parts of the population was developed. At the initial stage of the study, when data are collected, descriptive statistics and visualization methods will be used to identify primary dependencies, taking into account the type of the input data scale.

Survival analysis, also known as event history analysis, is an improved statistical method that is used to assess the probability of an event that occurs over time.

Results: The article presents the design of the framework for the study of the turnover readiness of economically active parts of the population. The implementation of methods of intellectual analysis based on survey data will help to form a workable approach to forecasting of the labor market needs.

Conclusion: Data Mining is one of the new technologies in the field of data analysis tools, and has a significant impact in the field of human resources. The corresponding framework, according to the initial data collection, will provide a qualitative set for various methods Data Mining.

Надійшла до редколегії 25.09.2017