26. Sobolevskyi, R. Cluster analysis of fracturing in the deposits of decorative stone for the optimization of the process of quality control of block raw material [Text] / R. Sobolevskyi, N. Zuievska, V. Korobiichuk, O. Tolkach, V. Kotenko // Eastern-European Journal of Enterprise Technologies. – 2016. – Vol. 5, Issue 3 (83). – P. 21–29. doi: 10.15587/1729-4061.2016.80652

27. Korobiichuk, V. Definition of hue of different types of pokostivskiy granodiorite using digital image processing [Text] / V. Korobiichuk, V. Shamrai, O. Iziumova, O. Tolkach, R. Sobolevskyi // Eastern-European Journal of Enterprise Technologies. – 2016. – Vol. 4, Issue 5 (82). – P. 52–57. doi: 10.15587/1729-4061.2016.74849

28. Vashchuk, O. M. Heostatystychnyy pidrakhunok zapasiv Velyko-Hadomynets'koho rodovyshcha pervynnykh kaoliniv z vrakhuvannyam sortovoyi dyferentsiatsiyi [Text] / O. M. Vashchuk, R. V. Sobolevs'kyy // Visnyk ZhDTU. Seriya: Tekhnichni nauky. – 2014. – Issue 1 (68). – P. 124–132.

*Розглядається комбінований підхід до прогнозування намірів співробітника підприємства звільнитися на основі демографічних факторів і факторів задоволеності роботою. Розроблений метод включає попередню оцінку надійності даних опитування співробітників, аналіз кореляційної залежності, побудову регресійної моделі і нелінійного предіктору плинності персоналу. Цей метод дозволяє менеджеру з управління персоналом не тільки виявляти співробітників, що потрапляють в зону можливої плинності, а також коригувати процеси управління персоналом базуючись на найбільш критичних факторах*

*Ключові слова: прогнозування, логістична регресія, управління персоналом, плинність персоналу, задоволення від роботи*

*Рассматривается комбинированный подход к прогнозированию намерения уволиться сотрудников предприятия на основе демографических факторов и факторов удовлетворенности работы. Разрабатываемый метод включает предварительную оценку надежности данных опроса сотрудников, анализ корреляционной зависимости, построение регрессионной модели и нелинейного предиктора текучки персонала. Этот метод позволяет менеджеру по управлению персоналом не только выявлять сотрудников, попадающих в зону возможной текучки, а также корректировать процессы управления персоналом, базируясь на наиболее критичных факторах*

*Ключевые слова: прогнозирование, логистическая регрессия, управление персоналом, текучесть персонала, удовлетворение от работы*

# INFLUENCE OF DEMOGRAPHIC FACTORS AND FACTORS OF JOB SATISFACTION IN THE PROCESSES OF PERSONNEL MANAGEMENT: PREDICTION OF STAFF TURNOVER BASED ON LOGISTIC REGRESSION

**N. Manakova**
PhD, Associate Professor, Head of Department*
E-mail: NatalyM@rambler.ru
**I. Tsyhanenko**
Postgraduate student*
E-mail: innatsyg@gmail.com
**G. Bielcheva**
PhD
Department of Media Systems and Technologies
Kharkiv National University of Radio Electronics
Nauka ave., 14, Kharkiv, Ukraine, 61166
E-mail: annijja@gmail.com
**O. Shtelma**
Senior Lecturer*
E-mail: olga.shtelma@gmail.com
*Department of Applied Mathematics and
Information Technologies
O. M. Beketov National University of
urban economy in Kharkiv
Marshala Bazhanova str., 17, Kharkiv, Ukraine, 61002

## 1. Introduction

In today's world, human factor role in the enterprise activity is difficult to overestimate. Planning in the field of personnel management is an integral part of strategic and tactical planning of company development. Tasks, which are set for planning in the personnel management area, should provide a company with human resources within a set time

limit and clearly plan organization of the process of effective staff hiring and development. Deviations from the plan cause significant unpredictable financial and time costs.

Every enterprise has its own peculiar internal structure, which is why a model of predicting an employee's decision-making in each individual case will include its own variety of factors.

When searching for ways to improve effectiveness of enterprise functioning under current conditions, the emphasis should be placed on a particular individual. It is clear that successful existence of any enterprise depends on the way the tasks are performed at each level of management and execution.

The most important task for human resource managers is the motivation creation for employees so that they should work better and more. However, motivation is related to the mental condition of a person. It forms both behavior and behavioral boundary, the space a person regards as his own, understanding of oneself and inherent personality.

Having a prediction assessment of an employee's decision to quit, an expert in the company's human resource department will be able to receive a signal, and pay attention to a particular employee or a group of employees. If employees belong to a risk group, it is necessary to respond timely with a system of measures for personnel policy.

Thus, prediction of behavior of an employee considering a change of work is the most important task of personnel management. In addition to explicit motivating factors that influence decision making of an employee, it is necessary to consider a number of demographic factors. The impact of these factors remains insufficiently studied, which leads to the staff turnover of a company and subsequent unforeseen financial and time costs. Therefore, study of methods to predict staff turnover is a relevant task.

## 2. Literature review and problem statement

A subject of effective human resource management has already become one of the central in the literature on personnel management. Identification of risks associated with personnel management has been actively covered in the literature for more than two decades [1]. However, the risks associated with human resources have not been explored at such a scale as other types of risks. This trend is predetermined by a variety of factors, complexity of their agreement, as well as a need to keep account of competence for specific jobs and positions. Paper [2] considered the capabilities of software, mathematical and statistical methods for the estimation and analysis of competence in personnel information systems. The use of quantitative analysis provides an accurate account of the competence, which may be used in many different areas. The standard model of competence, considering professional, innovative and social factors was presented. However, one should note its inertness in terms of risks of staff turnover.

Thus, in [3], authors considered articles related to risks in the field of personnel management. The studies focus on a review of more than 150 publications, since 2000 and up to now, which considered human resource problems through the prism of risks for a company. In [3], the need for more comprehensive and large-scale research into personnel resources was substantiated.

A variety of models of human resource management explores a relation between the policy of a human resource department and company's management and the degree of job satisfaction and risk of employees' quitting [4]. Let us consider these factors in more detail. In particular, article [5] presents a whole variety of psychometric information. Such components of the personality profile as energy, commitment, professional enthusiasm were distinguished. The proposed aspects. in combination with professional qualities, provide the largest efficiency and reduce the risk of employees' quitting. It is also possible to follow the link between the risks related to social activity and communicative traits of a personality [6]. At the same time, work satisfaction depends on demographic characteristics [7]. Paper [8] highlighted a number of the following aspects: possibility of professional growth, the size of an enterprise, salary level. Psychological state of an employee also depends on stressful situations that can be caused by a variety of tasks or requirements for new or different skills of an employee [9]. Comfortable working conditions are often associated with the lack of psychological pressure in a team [10]. It should be noted that the assessment of a degree of job satisfaction is more objective in the case of using surveys when an employee can independently define his emotions and mood [11].

The study of approaches to the prediction of staff turnover was paid particular attention to throughout the whole period of existence of human resource management. However, most studies still focused on labor productivity, neglecting the psychometry of an employee. In article [12], authors successfully apply some aspects of a personality profile. To forecast the prospects for a successful career in a company, such factors as a degree of employee's emotional resilience to load and stress were highlighted. In paper [13], on the basis of logistic regression, results of the forecasting model regarding relation between controlling variables and staff turnover were confirmed.

Given above studies [1–13], and taking into account a variety of indicators that affect staff turnover, the use of full automation for building a predictor will lead to the loss of information. For the purpose of objective risk assessment, it is necessary to engage specialists from a human resource department at different stages of prediction. This way of modeling the predictor on the basis of logistic regression will make it possible to take into account current human resource strategy of a company.

That is why in the present study main focus is aimed at the development of methods for predicting behavior of employees and motivational and demographic factors that influence decision making on changing a job of an employee.

## 3. The aim and objectives of the study

The aim of present study is to predict behavior of an employee concerning making a decision on changing a job with regard to demographic factors and factors of job satisfaction.

To accomplish the set aim, the following tasks had to be solved:

– to explore certain processes in the management of staff turnover in order to detect existing data sets and possible risk factors for an increase in turnover;

– to develop a method for predicting the intention to quit and for the identification of the most critical factors that will

take into account basic demographic features and factors of job satisfaction;

– to perform experimental verification of the developed method at all stages from data preprocessing to modeling the predictor of the intention of an employee to quit his position.

## 4. Methodology for the prediction of staff turnover

### 4. 1. Collection and preparation of data

For the purpose of analytical study, and in order to model staff turnover, it is planned to examine the following basic datasets that usually exist (or may be collected) at an enterprise:

1. Personal data collected on the basis of personal records of Human Resource Department. This dataset usually includes basic demographic data such as age, gender, marital status, existence of minor children and others. Additionally, the data, which describe the general indicators of staff turnover in separate spheres of activity of a company or at some positions, may be taken into account. It is also necessary to consider the reasons for previous quitting, the number of years that an employee spent in various positions, and others.

2. Results of employees' surveys on their degree of satisfaction with the current place of employment. Such questionnaires usually include questions that characterize self-realization in the profession, salary level, relations in the collective, respect in society and others. One factor for making an important decision as for one's career prospect is not, as a rule, enough.

Most often, when answering the questions, an employee should express the extent of his agreement/disagreement by the Likert scale, which may take the following form:

– 1 = do not agree at all;
– 2 = do not agree;
– 3 = rather do not agree;
– 4 = rather agree;
– 5 = agree;
– 6 = totally agree.

3. The last but the most indispensable element of this set is the data whether an employee quitted or did not, after a control period of time (quarter, half a year, year), after conducting a survey.

On the basis of such aspects, a specialist in charge of development and management of human resources will be able to respond timely and pay attention to a particular person. For example, to suggest taking qualification-upgrading courses, or changing in funding, etc.

Applying a policy of preventive steps, a company will get a real opportunity to decrease losses as a result of quitting of qualified personnel.

A number of factors that need to be analyzed in order to identify the employees that fall into a risk zone of quitting may reach several dozen. Multi-dimensional analysis cannot be performed only on the basis of the expert (human) assessment. Automated construction of full-scale models also does not meet testing prediction accuracy. In addition, with the use of automated approach, expert opinion is completely ignored, which is also a disadvantage. The best solution would be to develop a combined automated approach that will make it possible to obtain mathematical evaluation and to engage experts at every stage of selection of the most significant factors.

### 4. 2. Method of modeling a nonlinear predictor of staff turnover

The method of modeling a nonlinear predictor, which is being developed, includes 4 main stages to examine an influence of demographic factors and factors of job satisfaction on the staff turnover. The study is carried out from data agreement through evaluation of correlation links to the construction of a regression model. At each stage, it is possible to repeat several times a procedure for selection of the most optimal set of factors, relevant to the experience of an expert and personnel policy of the company. At the last stage, modeling of the predictor by the computed coefficients of the regression model is performed.

Let us enumerate the stages of modeling the nonlinear predictor:

1. Conducting interviews and testing coherence of the obtained responses.

2. Analysis of correlation between characterizing factors and staff turnover.

3. Creation of a generalized regression model of dependence of turnover on selected factors.

4. Modeling of nonlinear predictor on the basis of logistic regression by the coefficients of the model, constructed at the previous stage.

*Stage 1.* Testing consistency of survey data and making decision on their combination.

To assess internal consistency (reliability) of testing results, it is supposed to use Cronbach's Alpha as a degree of consistency, known also as Guttman's $\lambda_3$.

Cronbach's Alpha is calculated by comparing a rating for each element of the questionnaire with a general summary result for each respondent and successive comparison with the variances of individual elements:

$$\propto = \frac{k}{k-1}\left(1 - \frac{\sum_{i=1}^{k}\sigma_{Y_i}^2}{\sigma_X^2}\right), \tag{1}$$

where k is the number of elements in the questionnaire, i is the number of respondents, $Y_i$ is the answers of the i-th respondent, X is the total result of every respondent, $\sigma_{Y_i}^2$ is variance of the i-th element, $\sigma_X^2$ is variance of the final result.

The resulting coefficient of reliability $\alpha$ may take values from 0 to 1. If all elements of the questionnaire (or its subgroup) are completely independent on each other (that is, not correlated), then $\alpha=0$; and, conversely, if all elements have high co-variations, then $\alpha$ is approaching 1. In other words, the higher coefficient $\alpha$, the higher total co-variance of the elements and the higher the probability that all measurements are parts of one and the same basic concept.

Although the standards on which factor $\alpha$ is "good" are completely arbitrary and depend on your theoretical knowledge about the examined phenomenon, most methodologists recommend a minimum coefficient $\alpha$ between 0.65 and 0.8 (or higher in many cases). High consistency indicators allow data aggregation into groups, thus reducing the number of factors.

Coefficients $\alpha$, which are lower than 0.5, are usually unacceptable, especially for the questionnaires, which are supposed to be parts of one examined phenomenon. In this case, it is advisable to use elements of the questionnaire as separate, not interdependent factors.

It should be noted that in a number of cases responses to the survey would require preliminary scaling and rever-

sal. This is caused by the fact that some questions assume answers on scales with a varying number of points (for example, some by the agreement scale from 1 to 6 (as in our example) or from 1 to 10, some on the frequency scale from 1 to 5 (frequently, regularly, occasionally, sometimes, never)). Quite often questions are compiled not only in a positive manner, but also in negative, requiring an appropriate reversal before the test.

*Stage 2*. Correlation analysis between factors and staff turnover.

The method for calculation of correlation coefficient depends on the type of scale that includes variables. Thus, to measure variables with interval or quantitative scales, it is necessary to use the Pearson correlation coefficient. If at least one of the two variables has the ordinal scale, or is not normally distributed, it is necessary to use the Spearman ranging correlation or Kendall tau.

In the case of socio-demographic and personnel data, several different types of scales, including ordinal scale (survey data by the Likert scale), rational scales (for example, age), nominal scales (for example, reasons for quitting previous job) are used.

In addition, the dependent variable is represented in the binary (dichotomous) scale. In the case where one of the variables is dichotomous, it is recommended to use a rank-biserial correlation. This coefficient is a special case of Somers ' D.

Rank-biserial correlation is calculated by the following formula:

$$R_{rb}=2\times(Y_1-Y_0)/n,$$

where n is the number of data pairs in a sample, $Y_0$ is the mean value Y for data pairs with x=0, $Y_1$ is the value of Y for data pairs at x=1.

After calculation of correlation coefficients, a researcher may exclude a part of the factors from subsequent processing in the case their quantity is large. Excessive number of factors typically decreases the quality of a regression model.

*Stage 3*. Construction of a regression model with selected factors and staff turnover outcome.

The problem of prediction of a certain value or a specific object behavior by the set of independent factors can be solved using a regression analysis. In a general form, a mathematical model, which is constructed on the basis of experimental data, takes the following form:

$$Y=f(x_1, x_2,...,x_i,... ,x_n )+\varepsilon,$$

where f is the dependence function (the analytical form of which is examined); $x_1$, $x_2$,..., $x_i$,..., $x_n$ are the vector of values of influencing factors (or variables that explain a result); Y is the result or response; $\varepsilon$ is the random error.

The analytical method choice is based on several positions:
– number of influencing factors;
– type of explanatory data;
– type of resulting variable;
– probable laws of distribution of explanatory data and the resulting variable.

Because the examined dependent variable is dichotomous (binary), the appropriate method of regression analysis is logistic regression. Logistic regression is used to describe data and to explain the relationship between one dependent binary variable and one or more nominal, ordered, interval or rational quantitative independent variables.

Mathematically, logistic regression assesses a set function of linear regression, which is determined as follows:

Let $x=(x_1,..., x_p)^T$ be a vector of explanatory variables, and the output variable takes y=1 or y=0. Then, the model that describes the output variable takes the following form

$$\text{logit}(p_i)=\ln\left(\frac{p_i}{1-p_i}\right)=\beta_0+\beta_1 x_{1,i}+...+\beta_m x_{m,i}.$$

When constructing a regression model in logistic regression, one of the most important concerns is overfitting of the model. Addition of independent variables to the logistic regression model will always increase its statistical reliability. However, though model extension by increasing number of variables improves its data consistency, this way substantially impairs prognostic characteristics of the model. Therefore, the first two stages of the developed method allow a researcher to decrease the number of influencing factors. As a result of the model construction and after quality evaluation tests conducting, its coefficients will be computed, as well as an assessment of which of the influencing factors are significant and which are not. To improve the model, it is sometimes advisable to perform sequential reduction of influencing factors under an information criteria control, like the Akaike information criterion (AIC).

*Stage 4*. Construction of a nonlinear predictor of staff turnover.

Logistic regression implies that the dependent variable is a random event, and, in this sense, the constructed logistic regression may be used as a prediction function. We model conditional probability y=1, designated as p (y=1|x)=π(x), using the previously constructed model of logistic regression. Then the non-linear predictor of event occurrence Y=1 will take the following form:

$$P(Y=1)=\frac{1}{1+e^{-(\beta_0+\beta_1 x_{1,i}+...+\beta_m x_{m,i})}}.$$

To check the quality of modeling the predictor, it is advisable to split the existing dataset into a training sample (for direct construction of the model at the third stage) and a reference sample (for checking the quality of prediction at the fourth stage).

Having applied the constructed predictor to the existing staff structure, it is possible to identify employees who belong to the group with turnover risk, and thus require special attention from the human resource department and the company top managers.

## 5. Verification of developed methods and models

For experimental verification of the developed method for analysis, a sample of 216 respondents at 17 variables was prepared: 8 demographic and personnel factors, 8 satisfaction factors and an output dichotomous variable "quit/did not quit". All factors (name and content) are given in Table 1.

Job satisfaction questionnaire included eight questions:
a. If I want to, I have real opportunities to be promoted;
b. My work place is conveniently located from my home;
c. It is interesting for me to solve problems that arise in my work;

d. I do not like the way my boss treats me;

e. I have a convenient work schedule;

f. I feel pride and sense of personal accomplishment from my work;

g. I do not get satisfaction from working with their colleagues;

h. I receive appropriate payment for the work I perform.

The respondents gave all the answers by the same six-point Likert scale (absolutely disagree; disagree; rather disagree; rather agree; totally agree). That is why the pre-scaling of these data is not required. Most of the questions have a positively oriented scale (a, b, c, e, f, h), but some of the questions have a negatively oriented scale (d, g). Therefore, the responses received to these questions should be reversed.

The developed method was experimentally implemented in the environment RStudio using the program language of statistical computations R. At different stages, the following packages were used: psych (the function alpha() to calculate the Cronbach's Alpha, the function biserial() to calculate correlation coefficient using the method of ranking biserial correlation), and the packet glm (the function glm() to construct a regression model and function predict() to calculate prediction values).

Preliminary review of data and elementary calculations using the functions summary() and mean() from the programming language of statistical computations R allowed us to observe the following:

1. Statistics of the staff turnover according to column Quit. The number of employees who quit makes up only 15 % (32 employees from 216), the majority – respectively 85 % (124 employees) continued working.

2. Mean values by all the job satisfaction factors are larger than the a priori mean value that is equal to 3.5 for all scales since one to six.

Mean values for all factors are given in Table 2. Thus, it may be concluded that most employees are satisfied with their job.

At the first stage of the research, evaluation of reliability of the data on the results of the survey is conducted, that is, answers from 216 respondents to 8 questions concerning job satisfaction factors. Calculation of the Cronbach's Alpha by formula (1) and using the function psych::alpha() in R showed that the consistency coefficient of job satisfaction factors makes up 0.66.

Table 1

### Formalized designations of factors that influence decision making

| Name of indicator | Name of variable for calculations | Content of indicator |
|---|---|---|
| Demographic indicators | | |
| Gender | Gender | Gender of employee. *Indicator is measured on nominal scale.* |
| Age | Age | Age of employee. *Indicator is measured on rational scale.* |
| Marital status | Marriage | Indicator reflects current marital status of an employee. It is measured in points from 0 and 1, where unity indicates that an employee is married, zero that he is single. *Indicator is measured on dichotomous scale.* |
| Children | Children | The number of minor children. *Indicator is measured on ordered scale.* |
| Education | Education | Education of employee. It is measured in points, where 1 is higher education, 2 is special education, 3 is secondary education. *Indicator is measured on nominal scale.* |
| Personnel factors | | |
| Average term of employment | YearStay | Average term an employee has worked in one position is measured in years. *Indicator is measured on pseudometric scale.* |
| Reason of changing previous job | ChangeReason | This *indicator is nominal* and is measured from one to seven points: 1 is dissatisfaction with salary level, 2 is a conflict, 3 is making position redundant, 4 is enterprise closure, 5 is change of residence, 6 is change of working conditions, 7 is for personal reasons. |
| Turnover in professional area | PercentQuit | Turnover in professional area is represented in percentage form. *Scale of indicator is rational.* |
| Job satisfaction (by a survey). All scales of indicators in this section include from 1 to 6 – ordered Likert scale | | |
| Payment rate | Pay | I get appropriate payment for the work I perform. |
| Atmosphere in workforce | Colleagues | I do not enjoy working with my colleagues (reversed for further analysis) |
| Self-realization in profession | Realization | I feel pride and sense of self-realization from my work |
| Relationships with authorities | Chief | I do not like the way my boss treats me (reversed for further analysis) |
| Interesting tasks | Interest | It is interesting for me to solve problems that arise in my work |
| Career growth | Career | If I want to, I have a real opportunity to be promoted |
| Convenience of work site location | Situated | My place of work is conveniently located from my home |
| Working schedule | Scheduled | I have convenient working schedule |
| Output dichotomous variable | | |
| Quit | Quit | Indicator whether employee quits in a year (did =1, no=0) |

Table 2

Mean values of job satisfaction factors

| Number of factor | Job satisfaction factors | | Mean |
|---|---|---|---|
| 1 | Salary rate | Pay | 4.389 |
| 2 | Atmosphere in workforce | Colleagues | 4.012 |
| 3 | Self-realization in profession | Realization | 4.354 |
| 4 | Relationships with management | Chief | 4.305 |
| 5 | Challenges | Interest | 4.68 |
| 6 | Career growth | Career | 4.74 |
| 7 | Conveniently located workplace | Situated | 4.268 |
| 8 | Working schedule | Scheduled | 4.56 |

This indicator shows that reliability (consistency) of answers is acceptable, but different phenomena are possibly described. At this point, an expert on human resource management may want to consider a possibility of combining the indicators into groups (to reduce the number of variables in the model) or to consider them as separate factors. Table 3 below shows indicators of the Cronbach's Alpha in the case of combining the factors into three groups:

Table 3

Indicators of consistency of survey factors

| Group and name of variable | Factors included in group | α |
|---|---|---|
| Motivating work Motivation | Self-realization Challenges Career growth | 0.72 |
| Acceptable conditions Conditions | Appropriate salary Convenient location Convenient work schedule | 0.72 |
| Good team Team | Relations with management Atmosphere in team | 0.84 |

Upon obtaining these indicators of the Cronbach's Alpha, the expert has to decide whether to combine the factors. According to the indicators, this is possible, but not necessary. Let us assume that a decision was made on the aggregation of the agreed factors. Thus, the job satisfaction factors list will be reduced from eight to three.

The second stage of the study is focused on the factors identification of factors that influence employee choice whether to continue to work or quit. Let us verify the hypothesis on that not a single factor impacts an employee's decision whether to stay or quit.

To verify the hypothesis, if the relationship between the factors and the resulting variable is significant, the correlation coefficient as relationship indicator was measured (Table 4).

The table uses a designation of a relation degree: *** is the strongest, ** is strong; * is weak. Alternative hypothesis: an employee's decision whether to change the place of work depends on a specific factor (or factors). As stated above, based on the nature of data, to determine a correlation, a method of ranking bi-serial correlation is used.

According to the results of correlation coefficient, an alternative hypothesis is accepted. So could be reduced the following factors: Age, Education, Average term of work, Reason for change of previous job and Turnover in professional area. Thus, after the second stage, 7 factors remained.

Passing to the third phase, in order to obtain quality evaluations of predictive values, all existing dataset was divided into two samples (test and control). Using the t-Student criterion, both samples were checked for a difference in mean magnitudes and it was concluded that the difference between groups is statistically insignificant, which is why these two samples may be used for analysis.

Table 4

Calculation of correlation coefficient

| Factor | Correlation coefficient |
|---|---|
| Gender | 0.52** |
| Age | 0.32* |
| Marital status | −0.38* |
| Number of children | −0.48* |
| Education (1 – higher/ 2 – special/ 3 – secondary) | 0.07 |
| Average term of working in one position, years. | 0.09 |
| Reason for changing previous place of work | 0.103 |
| Turnover in professional area, % | −0.02 |
| Motivating job | −0.523** |
| Acceptable conditions | −0.783*** |
| Good team | −0.37* |

The first variant of the regression model will be constructed in the form of:

$$Y = \log_e(\frac{\overline{p}}{1-p}) = \beta_0 + \beta_1 * Age +$$
$$+\beta_2 * Gender + \beta_3 * Marriage +$$
$$+\beta_4 * Conditions + \beta_5 * Motivation + \beta_6 * Team.$$

Obtaining result from a logistic regression (Fig. 1), it was found that factors Good team (variable Team in Figure) and Age do not make a significant contribution to the model and may be excluded.

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.777308   0.313515   8.859 3.08e-13 ***
Gender      -0.127823   0.069460  -1.840   0.0697 .
Age          0.003190   0.005962   0.535   0.5943
Marriage    -0.149078   0.077526  -1.923   0.0583 .
Children    -0.212646   0.039465  -5.388 8.10e-07 ***
Conditions  -0.341134   0.034778  -9.809 5.01e-15 ***
Motivation  -0.092882   0.036477  -2.546   0.0130 *
Team        -0.015739   0.025627  -0.614   0.5410
```

Fig. 1. Modeling of a logistic regression

In the reduced model, shown in Fig. 5, each factor is statistically significant. Thus, there are 3 demographic indicators and 2 job satisfaction indicators left.

$$Y = \log_e(\frac{\overline{p}}{1-p}) = \beta_0 + \beta_1 * Gender +$$
$$+\beta_2 * Marriage + \beta_3 * Children +$$
$$+\beta_4 * Conditions + \beta_5 * Motivation.$$

Fig. 2 shows coefficients of calculation of the reduced regression model.

It should be noted that, according to the results, the Intercept has the highest absolute value. This indicates that the model has a latent factor(s), which was not revealed in the present study. Probably, the expert has to return to the previous stages for reconsidering factors not included before.

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.85391    0.24882  11.470  < 2e-16 ***
Gender      -0.13211    0.06857  -1.927  0.05777 .
Marriage    -0.15677    0.07611  -2.060  0.04285 *
Children    -0.20809    0.03862  -5.388 7.72e-07 ***
Conditions  -0.34139    0.03442  -9.918 2.37e-15 ***
Motivation  -0.09813    0.03427  -2.864  0.00541 **
```

Fig. 2. Construction of the reduced model of logistic regression

In addition, all the model coefficients (except the Intercept) are negative, which corresponds to the fact that the output variable takes the value of 1 in a negative case, that is, when an employee quits. In other words, the higher the relevant factors, the lower the probability that an employee will have the intention to quit the job.

Table 5

### Values of key coefficients

| Variable | Factor | Coefficient of model |
|---|---|---|
| Gender | Gender | −0.13 |
| Marriage | Marital status | −0.157 |
| Children | Existence of children | −0.208 |
| Conditions | Acceptable conditions | −0.34 |
| Motivation | Motivating job | −0.1 |
| Intercept | Free coefficient | 2.85 |

Completing the third stage, regression model is constructed:

$$Y = \log_e\left(\frac{\bar{p}}{1-\bar{p}}\right) = 2{,}85 - 0{,}13 * \text{Gender} - 0{,}157 * \text{Marriage} - 0{,}208 * \text{Children} - 0{,}34 * \text{Conditions} - 0{,}1 * \text{Motivation}.$$

At the fourth stage, a nonlinear predictor will be modeled as presented below (2) and the prediction result if an employee from a control sample quit the job will be checked.

$$P(Y=1) = \frac{1}{1 + e^{-(2{,}85 - 0{,}13 * \text{Gender} - 0{,}157 * \text{Marriage} - 0{,}208 * \text{Children} - 0{,}34 * \text{Conditions} - 0{,}1 * \text{Motivation})}}. \quad (2)$$

Using the modeled non-linear predictor, predictive values in the control group are obtained, the fragment of which is shown in Fig. 3.

| QuitProbs | QuitPredict | Quit |
|---|---|---|
| 0.46038761 | 0 | 1 |
| 0.52793424 | 1 | 1 |
| 0.64626807 | 1 | 1 |
| 0.13152965 | 0 | 0 |
| 0.42569727 | 0 | 1 |
| 0.90574535 | 1 | 1 |
| 0.71564876 | 1 | 1 |
| 0.02773260 | 0 | 0 |
| 0.23770875 | 0 | 0 |
| -0.33132559 | 0 | 0 |
| 0.06480500 | 0 | 0 |

Fig. 3. Predictive and existing values of quitting the job in a control group

Obtained results are interpreted as follows: with a probability from QuitProbs column we shall obtain the value of the output variable that is equal to unity, that is, in our case, an employee will quit. QuitPredict column is calculated according to the threshold value. If the probability exceeds 0.5, then the predictive value of quitting the job is 1, otherwise it is equal to 0. On the basis of these data, one calculates accuracy of prediction as the ratio of the means between the obtained (QuitPredict) and the known values (Quit), which for the present experimental set was 0.875.

## 6. Discussion of results and further directions of research

The developed method for the prediction of probable scenario of behavior of employees relative to the change of the place of work is distinguished by the possibility to determine the factors that affect staff turnover. Thus, it is possible to take into account individual factors and aspects of job satisfaction in the context of specific activity of a company.

This method includes preliminary assessment of reliability of the data of staff survey, analysis of correlation dependence, construction of a regression model and a non-linear predictor to assess the probability of staff turnover. At the stages of the method, an employee of a company has a possibility to reduce the number of factors (by grouping or discarding insignificant factors). Such a choice is made both on the basis of mathematical indicators and taking into account experience of the expert from a human resource department. To keep an expert component, authors of the present study refused to apply the automated methods for reducing dimensionality, such as a Principal Component Analysis. However, such refusal could lead to a significant number of factors that would be rather cumbersome for expert processing by a human and might lead to overfitting of the prediction model. To eliminate such probable shortcomings, it is possible to replace the expert component with the introduction of weighted coefficients for the factors that agree with strategic personnel management of respective company and take into account experience of the experts in a relevant field.

In addition, introduction of the method requires conducting of preliminary (or regular) surveys and specific resources for their processing from a human resource department. Therefore, implementation of the developed method will be advisable among medium-sized and large companies in the state sector or big business. The use of the method is most appropriate in industries where the value of personnel is quite high due to insufficient number of specialists or lengthy periods of training skilled employees.

It should be noted further implementation of the method (in the work of a human resource department in a company) makes it possible to conduct surveys among employees of the company on a regular basis and to accumulate panel data on staff turnover.

Applying the results of implementation of the developed method and data accumulation by experts from human resource departments, it is planned to develop the method using an analysis of panel data in order to consider the dynamics of staff turnover and a reaction to the policy of preventive measures. In addition, it is planned to develop a weight criterion to take into account the value of individual employees and a substantiated reduction of threshold value that relates them to a risk zone.

## 7. Conclusions

1. We identified and analyzed existing sets of available data and possible risk factors of increased staff turnover. Among them, a group of distinct demographic features (such as gender, age, marital status, number of children) and job satisfaction factors (for example, salary level, atmosphere in the team, self-realization in the profession) is entered.

2. A method for the prediction of intention to quit the job was developed, which allows us to identify the most critical factors for staff turnover. The proposed method for modeling the non-linear predictor of staff turnover allows an expert from a human resource department to vary (to group or reject insignificant) factors taking into account the current strategy for company personnel management.

3. The developed approach was realized in the process of human resource management by practical verification. Based on a logistic regression, a group of demographic factors was revealed, such as gender, marital status and birth of children, as well as two aspects of job satisfaction: good working conditions and adequate motivation, which have an extremely negative impact on staff turnover. In addition, it was found that salary and marital status are significant predictors of the intentions for staff turnover. Thus, the developed method was implemented in studies into dynamics of staff turnover, from data preprocessing to modeling the predictor of the intention of an employee regarding quitting the job.

## References

1. Voegtlin, C. Corporate social responsibility and human resource management: A systematic review and conceptual analysis [Text] / C. Voegtlin, M. Greenwood // Human Resource Management Review. – 2016. – Vol. 26, Issue 3. – P. 181–197. doi: 10.1016/j.hrmr.2015.12.003

2. Bohlouli, M. Competence assessment as an expert system for human resource management: A mathematical approach [Text] / M. Bohlouli, N. Mittas, G. Kakarontzas, T. Theodosiou, L. Angelis, M. Fathi // Expert Systems with Applications. – 2017. – Vol. 70. – P. 83–102. doi: 10.1016/j.eswa.2016.10.046

3. Lannoo, S. What drives the drivers? Predicting turnover intentions in the Belgian bus and coach industry [Text] / S. Lannoo, E. Verhofstadt // Transportation Research Part A: Policy and Practice. – 2016. – Vol. 91. – P. 251–259. doi: 10.1016/j.tra.2016.06.024

4. Becker, K. A risk perspective on human resource management: A review and directions for future research [Text] / K. Becker, M. Smidt // Human Resource Management Review. – 2016. – Vol. 26, Issue 2. – P. 149–165. doi: 10.1016/j.hrmr.2015.12.001

5. Bakker, A. B. The Job Demands-Resources model: state of the art [Text] / A. B. Bakker, E. Demerouti // Journal of Managerial Psychology. – 2007. – Vol. 22, Issue 3. – P. 309–328. doi: 10.1108/02683940710733115

6. Shahpouri, S. Mediating role of work engagement in the relationship between job resources and personal resources with turnover intention among female nurses [Text] / S. Shahpouri, K. Namdari, A. Abedi // Applied Nursing Research. – 2016. – Vol. 30. – P. 216–221. doi: 10.1016/j.apnr.2015.10.008

7. Osman, I. The Dynamic Role of Social Exchange and Personality in Predicting Turnover Intentions among Professional Workers [Text] / I. Osman, F. Noordin, N. Daud, M. Z. Othman // Procedia Economics and Finance. – 2016. – Vol. 35. – P. 541–552. doi: 10.1016/s2212-5671(16)00067-8

8. Daud, N. Determinants of Job Satisfaction: How Satisfied are the New Generation Employees in Malaysia? [Text] / N. Daud // Procedia – Social and Behavioral Sciences. – 2016. – Vol. 219. – P. 208–213. doi: 10.1016/j.sbspro.2016.05.007

9. Zhao, X. (R.) Motivating frontline employees: Role of job characteristics in work and life satisfaction [Text] / X. (R.) Zhao, R. Ghiselli, R. Law, J. Ma // Journal of Hospitality and Tourism Management. – 2016. – Vol. 27. – P. 27–38. doi: 10.1016/j.jhtm.2016.01.010

10. Mathieu, C. Corporate psychopathy and abusive supervision: Their influence on employees' job satisfaction and turnover intentions [Text] / C. Mathieu, P. Babiak // Personality and Individual Differences. – 2016. – Vol. 91. – P. 102–106. doi: 10.1016/j.paid.2015.12.002

11. Trivellas, P. Exploring Workplace Implications of Emotional Intelligence (WLEIS) in Hospitals: Job Satisfaction and Turnover Intentions [Text] / P. Trivellas, V. Gerogiannis, S. Svarna // Procedia – Social and Behavioral Sciences. – 2013. – Vol. 73. – P. 701–709. doi: 10.1016/j.sbspro.2013.02.108

12. Valle, M. A. Using self-organizing maps to model turnover of sales agents in a call center [Text] / M. A. Valle, G. A. Ruz, V. H. Masias // Applied Soft Computing. – 2017. doi: 10.1016/j.asoc.2017.03.011

13. Gloor, P. A. Forecasting managerial turnover through e-mail based social network analysis [Text] / P. A. Gloor, A. F. Colladon, F. Grippa, G. Giacomelli // Computers in Human Behavior. – 2017. – Vol. 71. – P. 343–352. doi: 10.1016/j.chb.2017.02.017