

UDC 004.8 (045)

DOI:10.18372/1990-5548.59.13634

V. M. Sineglazov

## INTELLIGENT SYSTEM OF PERSONNEL MANAGEMENT

Aviation Computer-Integrated Complexes Department, National Aviation University, Kyiv, Ukraine

E-mail: svm@nau.edu.ua ORCID 0000-0002-3297-9060

**Abstract**—It is considered the problem of Human Resources Management with help of Artificial Neural Networks in area of desired job search and help of recruiters that use this resource to find the best candidate for a given job. It is proposed to use neural networks for view CVs, ranking candidates according to their skill level and create machine learning algorithms to automate processes for checking resumes. The convolutional neural network is used for the problems solution. As a learning algorithms is used a genetic algorithm. It is developed an approach for chromosome structure optimal choice.

**Index Terms**—Human resource management; neural network; machine learning algorithm; genetic algorithm; chromosome structure; optimal choice.

## I. INTRODUCTION

Today, every company operating in the market can not do without the services of employees. Employees have played and continue to play a key role in the work of the enterprise, which poses a difficult task for managers to select highly qualified personnel. The solution of this problem is possible through the use of employee assessment methods and systems in the enterprise.

Personnel assessment is “management technology aimed at achieving the goals of the company and at implementing its strategy, as well as at increasing the effectiveness of the organization’s activities on the main management functions” [1, p. 30].

From the moment of the definition of personnel evaluation, various assessment methods began to emerge, some of which were automated. At the moment there are a lot of methods and automated personnel assessment systems used in different companies, but each method has its own characteristics, sphere of application advantages and disadvantages. It should be noted that each company is aware of the probability of using not the best methods and systems for evaluating employees.

Today, artificial intelligence (AI) is used as the main approach to solving human resource management and recruiting tasks.

Another application of AI is human resource management and recruiting. There are three ways to use AI for human resource management and hiring specialists. Artificial intelligence is used to view CVs and rank candidates according to their skill level. Artificial intelligence is also used to predict the success of a candidate in given roles through job matching platforms. Finally, AI is used to create chat bots that can automate repetitive communication tasks.

As a rule, the process of viewing a resume includes the analysis and search for information in the database of the resume. Startups, such as Pomato, create machine learning algorithms to automate processes for checking resumes. The Pomato AI system [1] is aimed at automating the verification of technical applicants for positions in technical firms. Artificial intelligence Pomato performs more than 200,000 calculations on each resume in seconds, and then develops its own technical interview based on useful skills.

From 2016 to 2017, Unilever consumer goods company used artificial intelligence to display all entry-level employees. Unilever AI used games based on neuroscience, recorded interviews and analysis of facial / speech signals to predict the success of a candidate in the company. Unilever collaborated with Pymetrics and HireVue to create a new AI-based analysis system and increase the number of candidates under consideration from 15,000 to 30,000 in one year. Unilever also reduced the time to process applications from 4 months to 4 weeks and saved over 50,000 hours of recruiters time.

From screening resumes to neuroscience, speech recognition and facial analysis, it is clear that AIs have a huge impact on the field of human resource management. The latest development in AI is the development of recruitment chats. TextRecruit, released Ari (automated recruiting interface). Ari is a recruitment chat room that is designed to conduct two-way text conversations with candidates. Ari automates the publication of vacancies, advertisements, screening of candidates, planning interviews and developing relations of candidates with the company as you move through the recruitment process. Ari is currently offered as part of the TextRecruit project participation platform.

This paper discusses the problem of using AI (neural networks) to view summaries and rank candidates according to their skill level.

## II. NATURAL LANGUAGE PROCESSING

Natural Language Processing, (NLP) is the general direction of artificial intelligence and mathematical linguistics. It studies the problems of computer analysis of natural languages. The main areas of natural language processing include such as extracting facts, analyzing the text tonality, answering questions, information retrieval, text generation, translation, etc. By extracting information is meant to search in an unstructured or weakly structured document of individual facts of interest to you.

Text tonality analysis implies an automatic determining the emotional coloring of the text and identifying the attitude of the person who wrote the text to the object of discussion.

To solve the above problems, researchers use a huge set of tools and techniques of natural language analysis.

Typically, the task of classifying texts is solved by performing the following steps:

- 1) Pre-processing of texts.
- 2) Transferring of texts into the real space of features, where each document is associated with a vector of fixed length.
- 3) The choice of machine learning algorithm for classification.

The first stage is to remove words that do not carry information about this text. In the Ukrainian language, they can include prepositions, conjunctions, particles. The second stage of preprocessing of texts is to bring each word to the basis, the same for all its grammatical forms. This is necessary because the words bearing the same meaning can be written in different forms.

Most modern machine learning algorithms are focused on the characteristic description of objects, so all documents are usually transferred to the real space of features. The most well-known methods for transferring text into feature space are based on statistical information about words: Bag of Words [7], Bag of Words & TF IDF [8], Bag of Ngrams & TF IDF [9].

When using these methods, each object is transferred into a vector, the length of which is equal to the number of words used in all the texts of the sample.

Problems of the traditional text classification method:

- 1) To implement this method, it is necessary to choose a method for transferring text into a vector

representation, since usually for various tasks the best quality of classification is shown by various methods.

- 2) The resulting feature space will have a high dimension, and at the same time will be strongly discharged.

3) Most often, to improve the quality of the classification, it is necessary to remove the stop words from the text. When you delete a different set of stop words, you get a different result of the algorithm.

- 4) To effectively use this method, you must use stemming (deleting endings, bringing the word to the base) or lemmatization (bringing the word to the initial form), since words with different declensions have the same meaning.

In this paper, convolutional neural networks [10] are used to process textual information. The main approaches to the use of convolutional neural networks for the task of text classification are: character-by-character approach [2] (using convolutional neural networks), an approach using word coding [3] (using convolutional neural networks) when each word in the text is associated with a vector of fixed length, then from the obtained vectors for each sample object a matrix is compiled, which, similarly to the images, is fed to the input of a convolutional neural network.

The most well-known methods for representing a word using a fixed-length vector are: one-hot encoding, Word2Vec [4], Continuous Bag of Words (CBOW), Skipgram (Skipgram model works slower, but usually it helps to achieve better text classification quality), technology GloVe [5] (allows for each word in text data to obtain the corresponding vector of fixed length using statistical information about this word in the data).

A convolutional neural network (CNN) with a character-based approach for text classification has been implemented.

## III. PROBLEM STATEMENT

Consider the structural-parametric synthesis of a convolutional neural network with a character-by-character approach for text classification.

Let  $X$  be a set of object descriptions,  $Y$  is a set of numbers (or names) of classes. There is an unknown target dependence – the map  $y^*: X \rightarrow Y$  whose values are known only on objects of a finite learning sample:

$$X^m = \{(x_1, y_1), \dots, (x_m, y_m)\},$$

where  $X^m$  the set of elements of the training sample is dimensional  $m$ .

You need to build an algorithm structural-parametric synthesis of a convolutional neural network for the implementation of a character-by-character approach to the problem of text classification.

#### IV. PROBLEM SOLUTION

##### *A. Determination of the Most Significant from the Viewpoint of the Efficiency of Parameters of the Collective Neural Network*

The determination of the most significant in terms of the effectiveness of the parameters of the CNN was carried out as a result of an experiment on CNN.

For the experiment, a database MNIST (database of samples of handwritten writing of digits) of was used.

Planning an experiment, adhered to the following.

1) Analyzes the number of layers of convolution ranging from 2 to 4.

2) The aggregation layers will be placed after the convolution layers and in front of the sweep layers, one aggregation layer is used for one, two or three convolution layers. In each layer of aggregation, a  $2 \times 2$  core is used as the optimal one. With a larger aggregation core size, too much information is lost. In the aggregation layers, the maximization function is used because it is simple and reliable.

3) Layers of sweep will be used the same as the layers of convolution, in an amount of from 2 to 4.

4) The size of the convolution kernel must be odd and not too large, so sizes are possible – 3, 5, 7.

5) The number of attribute cards can be from 4 to 64, it is also desirable to use an even number, and it is desirable that the number be a power of two to simplify the calculations.

The offset will be used the same for the horizontal and vertical, always one to simplify the network architecture. Next, we will consider an offset of  $2 \times 2$  or  $3 \times 3$  for the convolution kernel of size 5 and 7.

To convolve the image matrix at the edges of the matrix, it is used to fill in the insufficient fields of the image matrix with zeros without reducing the feature maps.

As the activation function, the function ReLU is used, and the classifier is the NEFCLASS.

Given the above, the experiment was performed by changing the following parameters:

1) The number of layers of convolution from 2 to 4.

2) The number of layers of aggregation from 1 to 4.

3) Three options are considered: one aggregating layer after one convolutional layer, one aggregating layer after two convolutional layers, one aggregative layer after three convolutional layers.

4) Layers of convolution (on each layer separately):

– the size of the core of the convolution of  $3 \times 3$ ,  $5 \times 5$ ,  $7 \times 7$ ;

– the number of feature cards 8, 16 and 32;

– offset value is a fixed parameter;

– used edge effect – a fixed parameter.

5) Layers of aggregation:

– the size of the aggregation core is a fixed parameter;

– the function of the aggregation core is a fixed parameter.

6) Classifier:

– classifier type – fixed parameter;

– number of layers from 2 to 4;

– the size of each layer is 64, 128, 256 and 512.

7) Extraction operations for each layer: – the percentage of withdrawal -fixing parameter.

Fixed parameters of the neural network:

1) Layers of convolution:

– offset value 1 for each layer;

– at the edges of the image matrix for convolution is used to fill with zeros, the size of the feature maps is equal to the size of the image matrix

– the activation function is used by the ReLU.

2) Layers of aggregation:

– the size of the aggregation core  $2 \times 2$ ;

– as an option of aggregation, it is taken maximization aggregation (maxpooling).

3) Only fully connected layers are used – the ReLU activation function.

4) The extraction operation is not used.

5) Initialization functions – normalizing initialization (Glorot).

6) Error functions – MSE.

7) Number of training periods 20.

On the basis of the analysis of the results of the combined CNN check on the test sample, one can determine the significant parameters for the CNN:

1) Number of coil layers.

2) Number of layers of aggregation.

3) Mutual placement of the layers of the convolution and aggregate layers.

4) Layers of the convolution (on each layer separately):

– the size of the core of the convolution (on each layer separately);

– number of signs cards (on each layer separately);

- the value of the displacement (on each layer separately);
  - parameter of boundary effect.
- 5) Layers of aggregation (on each layer separately):
- the size of the nucleus of aggregation;
  - the function of the nucleus of aggregation.
- 6) Allowed layers (on each layer separately):
- number of fully bonded layers;
  - the size of each layer;
  - type of classifier: auto-encoder.
- 7) Presence of an extract operation for each layer is the percentage of extraction and random function.

To optimize the structure and parameters of CNN, a genetic algorithm is used.

#### B. The structure of the chromosome

In our algorithm, each instance corresponds to 114 bits. They can be represented in the form of 6 genes – by the number of rollers and classifiers. In the first three genes (the convolution genes), seven fields can be distinguished:

- 1) Presence of a layer of 1 bit (at 0 in the first gene an instance falls into the bioreactor).
  - 2) The number of charts of signs is 7 bits – from 4 to 256.
  - 3) The kernel size is 3 bits – the kernel has a size from 3 to 8 inclusive.
  - 4) Offset 2 bits – from 1 to 4 inclusive, with a displacement of more than half plus 0.5 ( $n / 2 + 0.5$ ), the specimen enters the bioreactor. That is, the most possible bias of 4 with the size of the core 7 or 8.
  - 5) Dropout, layer of the exception. The gene has a size of 5 bits, possible exclusion values of 0% to 30%.
  - 6) Regional effects of a convolution, 1 bit – definition of reducing the size of the image when performing a convolutional operation.
  - 7) A layer of aggregation (pooling), 3 bits. There are five possible options: the presence of a layer of aggregation, the size of the aggregation of  $2 \times 2$  or  $3 \times 3$  and the function of aggregation – maximizing and averaging.
- In the last three genes, three fields can be distinguished:
- 1) Presence of a layer of 1 bit (at 0 in the first gene an instance falls into the bioreactor).
  - 2) The size of the full-length layer is 11 bits – the layer has a size from 0 to 1024 inclusive.
  - 3) Dropout, layer of the exception. The gene has a size of 5 bits, possible exclusion values of 0% to 30%.

## V. CONCLUSION

In order to detect non-formalized elements, a methodology for processing video images based on autoconfigurators and CNNs has been developed, in which the optimal way of determining CNN parameters based on the use of the genetic algorithm.

## REFERENCES

- [1] 30+ Talent Acquisition Technologies That Use #ArtificialIntelligence <https://workology.com/30-talent-acquisition-technologies-that-use-artificialintelligence/>
- [2] Xiang Zhang, Junbo Zhao, and Yann LeCun, "Character-level convolutional networks for text classification," in *Advances in Neural Information Processing Systems*. pp. 649–657, Feb. 2015.
- [3] Yoon Kim, "Convolutional neural networks for sentence classification," *IEMNLP*, pp. 1746–1751, Sep. 2014.
- [4] Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean, "Efficient estimation of word representations in vector space," *ICLR*, 2013. arXiv:1301.3781v3 [cs.CL] 7 Sep 2013.
- [5] Jeffrey Pennington, Richard Socher, and Christopher D. Manning, "Glove: Global vectors for word representation," *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 1532–1543, October 25–29, 2014, Doha, Qatar.
- [6] Yoshua Bengio Yann LeCun, Leon Bottou, and Patrick Haffner, "Gradient-based learning applied to document recognition," *Proc. of the IEEE November 1998*, pp. 1–46.
- [7] Zellig Harris, *Distributional structure*, Word, 1954.
- [8] Jones K. S., "A statistical interpretation of term specificity and its application in retrieval," *Journal of Documentation*. 1972.
- [9] Marc Damashek, *Gauging similarity with n-grams: Language-independent categorization of text*. Science, New Series, 1995.
- [10] V. M. Sineglazov and O. I. Chumachenko, "Structural and Parametric Synthesis of Packing Neural Networks," *Materials of the International Scientific and Practical Conference "Information Technologies and Computer Modeling"*, Ivano-Frankivsk–Yaremche, Ukraine (May 14–19, 2018), pp. 225–228. (in Ukrainian)

Received September 21, 2018

**Sineglazov Victor.** orcid.org/0000-0002-3297-9060

Doctor of Engineering Science. Professor. Head of the Department.

Aviation Computer-Integrated Complexes Department, National Aviation University, Kyiv, Ukraine.

Education: Kyiv Polytechnic Institute, Kyiv, Ukraine, (1973).

Research area: Air Navigation, Air Traffic Control, Identification of Complex Systems, Wind/Solar power plant.

Publications: more than 600 papers.

E-mail: svm@nau.edu.ua

**В. М. Синеглазов. Інтелектуальна система управління персоналом**

У даній роботі розглянуто проблему управління персоналом за допомогою штучних нейронних мереж в області пошуку бажаної роботи і допомоги рекрутерів, які використовують цей ресурс для пошуку кращого кандидата на дану роботу. Пропонується використовувати нейронні мережі для перегляду резюме, ранжирування кандидатів відповідно до їх рівнем кваліфікації та створення алгоритмів машинного навчання для автоматизації процесів перевірки резюме. Згортова нейронна мережа використовується для вирішення завдань. В якості алгоритму навчання використовується генетичний алгоритм. Розроблено підхід до оптимального вибору структури хромосом.

**Ключові слова:** управління людськими ресурсами; база знань; система підтримки рішень; нейронна мережа.

**Синеглазов Віктор Михайлович.** orcid.org/0000-0002-3297-9060

Доктор технічних наук. Професор. Зав. кафедри.

Кафедра авіаційних комп'ютерно-інтегрованих комплексів, Національний авіаційний університет, Київ, Україна.

Освіта: Київський політехнічний інститут, Київ, Україна, (1973).

Напрямок наукової діяльності: аеронавігація, управління повітряним рухом, ідентифікація складних систем, вітроенергетичні установки.

Кількість публікацій: більше 600 наукових робіт.

E-mail: svm@nau.edu.ua

**В. М. Синеглазов. Интеллектуальная система управления персоналом**

В данной работе рассмотрена проблема управления персоналом с помощью искусственных нейронных сетей в области поиска желаемой работы и помощи рекрутеров, которые используют этот ресурс для поиска лучшего кандидата на данную работу. Предлагается использовать нейронные сети для просмотра резюме, ранжирования кандидатов в соответствии с их уровнем квалификации и создания алгоритмов машинного обучения для автоматизации процессов проверки резюме. Сверточная нейронная сеть используется для решения задач. В качестве алгоритма обучения используется генетический алгоритм. Разработан подход к оптимальному выбору структуры хромосом.

**Ключевые слова:** управление человеческими ресурсами; база знаний; система поддержки решений; нейронная сеть.

**Синеглазов Виктор Михайлович.** orcid.org/0000-0002-3297-9060

Доктор технических наук. Профессор. Зав. кафедры.

Кафедра авиационных компьютерно-интегрированных комплексов, Национальный авиационный университет, Киев, Украина.

Образование: Киевский политехнический институт, Киев, Украина, (1973).

Направление научной деятельности: аэронавигация, управление воздушным движением, идентификация сложных систем, ветроэнергетические установки.

Количество публикаций: более 600 научных работ.

E-mail: svm@nau.edu.ua