

GRANULAR CALCULATIONS IN THE REQUIREMENT ANALYSIS OF THE POLISH MARKET

Aneta Becker

Department of Economics
West Pomeranian University of Technology
e-mail: aneta.becker@zut.edu.pl

Jaroslav Becker

Department of Technology
The Jacob of Paradyż University of Applied Sciences in Gorzów Wielkopolski
e-mail: jbecker@pwsz.pl

Ryszard Budziński

Faculty of Economics and Management
University of Szczecin
e-mail: ryszard.budzinski@wneiz.pl

Abstract: The article presents the results of the analysis of the labour market requirements reported by Polish employers. Relations between the type of specialty, enhanced by the competence profile, and the requirement of proficiency in English were sought. The empirical material came from the research implemented within the “Human Capital Balance” project (V edition – 2014). The research procedure consisted of grouping objects with the method of *k*-means and the induction of decision rules based on the application of theories of rough sets. The analysis was performed using the RSES 2.1 system.

Keywords: competence, information granules, rough sets

INTRODUCTION

In the second half of the XX century L.A. Zadeh popularised the concept of the informative granule, also called the info-granule or information granule [Zadeh 1979, 1997]. The proposed term refers to certain sub-sets (classes) of the universe, which includes objects characterised by: indiscernibility, similarity or a similar behaviour. This term is associated with the formulation of granular calculations, which is defined by literature as the general theory of calculations. It aims

to effectively use the so-called grains, such as classes, clusters, sub-sets, groups and intervals. This will allow the construction of effective calculation models for purposes of complex applications with a significant amount of data, information and knowledge [Borowik et al. 2009, p. 154]. Granular calculations are perceived as the form of structural thinking or the procedure of structural problem solving. Depending on the assumed grain size (aggregation) in the tested set, one can observe the regularities (relations, dependencies) invisible with a different distribution, what allows a wider examination of the problem. The quality of results obtained from the application of the granular construction is affected by the structure of granules and methods of calculation implementation.

The aim of the article is to examine the competence needs of the Polish labour market. The main task is to analyse the competence demand (skills and qualifications) reported by the Polish employers. Studies were conducted in the paper using the RSES 2.1 system (Rough Set Exploration System 2.1), which is a computer tool, constructed by the team led by Z. Skowron. It enables the analysis of data in the table form using the rough set theory.

The studies used rough sets, used in the theory of granular calculations. The focus was on searching dependencies between the type of specialty, clarified by the competence profile, and the requirement of the English language proficiency. There were also attempts to answer three questions related to the analysed problem.

- Are there employers' requirements towards specialists concerning the knowledge of English?
- Which specialists (in particular) are required to know the English language?
- Is this regularity the same in the regions of Poland (information granules) characterised by a different level of economic development?

CONCEPT OF GRANULAR INFORMATION

Unlike traditional numerical calculations, data-oriented, granular calculations are directed to knowledge. That is why they are used in applications related to the discovery of knowledge and data exploration [Borowik et al. 2009, p. 155]. In 1982 Z. Pawlak proposed to use the granular information in the form of the rough set theory [Yao 2005].

The rough set theory helps to control large data sets and provides mathematical tools needed for a formal description of knowledge, especially the incomplete and inaccurate knowledge. Due to the fact that the granularity of the available information can cause inconsistency of the object description, this theory has been proposed as the tool of the granular information analysis [Nowak 2013]. The rough set theory allows, among others: searching dependencies between data, data reduction, determining the data weight, generating decision rules from data [Pawlak 1982]. In this method, the rule knowledge representation (in the form of IF... THEN) is the result of performing the decision rules induction process based on the set of data

representing the teaching models occurring in the object description. This description has the form of the decision table, where the rows represent next examples, and columns the selected features (describing attributes).

Among the methods of direct induction of decision rules it is possible to distinguish procedures based on the sequential coverage of the teaching cases set. This way we obtain the learning model, which has the form of a set of rules, which covers all cases from the learning set. An example of the algorithm of direct rule induction is LEM2 used in the studies [Grzymala-Busse 1992]. It is a popular procedure for creating a minimal set of rules, suitable for classification purposes. It creates a symbolic description discriminating the approximations of each decision class using the original proposal close to the rule of generating next covers. The LEM2 algorithm at the input gets the approximation of the analysed decision class. If this is the lower approximation, certain rules are generated, if the top one – the possible rules. In the study it is also possible to consider the marginal area, then the rough rules are generated [Skowron 1993].

CHARACTERISTICS OF WORK OFFERS FOR SPECIALISTS

The empirical material used in the conducted analyses came from studies implemented in 2014 within the “Human Capital Balance” project – HCB (V edition) [<http://bkl.parp.gov.pl/dane>]. It included the work offers (excluding internships and apprenticeships for students and apprentices) collected in the county labour offices (CLO) for all Polish provinces and at Careerjet.pl, a national web portal for job placement (job search engine). According to the research procedure, the offers were collected from the selected 160 CLO offices (10 offices per province). The sample was selected in an exhaustive manner, and the offers valid on the given day – on March 24, 2014. In case of the CLO the offer valid on the day of the conducted study was taken into account, while from the Careerjet.pl websites the first offers registered on this day were coded.

The article focused on information concerning offers directed to the following specialists:

- physical, mathematical and technical sciences (S1),
- for health matters (S2),
- teaching and education (S3),
- for economic and management matters (S4),
- for matters of the information and communication technologies (S5),
- from the field of law, social areas and culture (S6).

In the brackets there are designations of various groups of specialists adopted in the research.

Within the HCB studies the classification of competence included two groups of professional skills: formal (qualifications) and crucial. Formal competences related to: education – its level and profile (direction), experience (seniority) – the

course of the current employment, held certificates, permissions and diplomas, and additional resources (e.g. a car, Internet access, a computer). In contrast, crucial competences included skills: cognitive (searching and analysis of information and drawing conclusions), individual (psychological), cultural, physical (fitness), interpersonal (interpersonal contacts), management, disposition, office, technical (service, mounting and repairing devices), computer (computer skills and use of the Internet), mathematical (calculation), professional [Kocór et al. 2010].

In 2014 the most sought-after professional categories included: skilled workers, machine and equipment operators, specialists and sellers and service workers. In comparison to 2013 the labour demand has increased for: general practitioners, nurses, physiotherapists, professional subjects and language teachers, specialists for economy and programmers. In the case of skilled workers we sought: tillers, turners, mechanics, carpenters and seamstresses.

Among the requirements for candidates for future employees of utmost importance in 2014 was: experience (measured by the seniority at the similar offered position), level of the education and the skill to use a foreign language. The gender of candidates has slightly lost the significance. Employers recognised as the most importance competence: self-organisational, professional and interpersonal. It can be observed that there was a division of competence requirements related to the classification of professions into mental and physical. In the mental work the language, cognitive, computer and mathematical competence were more important. While in physical professions what counted was the physical fitness and technical abilities. On the basis of the owned information derived from candidates it was observed that they had insufficient professional experience and the inadequate motivation to work. Particularly clear was the gap regarding the competence: professional, self-organisational and interpersonal, that is the ones most valued by employers.

To regions with higher demand for specialists included: Lesser Poland and Silesian, Pomeranian, Opole, Świętokrzyskie, Warmia-Masuria and Podlasie. The specialists were most often sought to work by employers operating in the sectors of specialist services and services for the public (public and private education, health care and social care).

2014 was dominated by employers, who were searching for an employee, who could work without the long training. The more complex the obligations performed at work, the higher the expectations of employers within the professional experience. Also knowledge and formally certified skills, i.e. qualifications, gained on importance. Employers required formal documents confirming: practical abilities (e.g. a driving license, knowledge of resume writing, knowledge of specific computer programs), health care (e.g. current psycho-technical tests), taking part in obligatory training being the condition to undertake employment (e.g. safety training, fire training). For employers what also mattered was the fluent knowledge of English and German, most of all. It was important particularly for candidates for the so-called mental positions that is specialists, managers and office workers. While it was of no importance for physical professions.

ANALYSIS OF THE COMPETENCE NEEDS OF THE POLISH LABOUR MARKET

The study presented in the article concerning the search for dependencies between the type of specialty, clarified with the competence profile, and the requirement of the English language proficiency involved the use of the induction algorithm of the minimal set of decision rules (LEM2). The sample size used to perform calculations was 4636 offers. These offers came from particular regions of Poland. Data were included in the decision table, containing 12 conditional attributes, which took on the {yes/no} values and concerned the crucial competence: cognitive (k1), individual (k2), cultural (k3), physical (k4), interpersonal (k5), managerial (k6), disposition (k7), office (k8), technical (k9), computer (k10), mathematical (k11), professional (k12). The thirteenth attribute was related to the specialty –Spec{S1, S2, ..., S6}. While the knowledge of English {yes/no} was defined as a decision attribute.

In the first stage of the study the focus was on searching for the correctness in the whole group of offers directed to specialists (the main granule – 4636 offers). While in the second one the group was divided into 4 sub-granules, which were obtained as a result of classification of regions of Poland in terms of the wealth degree. The division into granules at this stage was taken from the work of [Becker 2014]. The empirical material included the selected macroeconomic categories and came from the CSO [Gross domestic product...]. The division was made using the method of *k*-mean [Grabiński 1992, p. 124-127], using the Statistica 10 program. The obtained sub-granules created the next classes of wealth of individual regions of Poland:

- class 1: Masovia, (sub1),
- class 2: Lower Silesia, Silesia, Greater Poland, (sub2),
- class 3: Kuyavian-Pomeranian, Lublin, Lodz, Lesser Poland, Pomeranian, West Pomeranian, (sub3),
- class 4: Lublin, Opole, Podlasie, Subcarpathian, Świętokrzyskie, Warmia-Masuria, (sub4).

Performing calculations the coherence of calculations was not studied, because the LEM2 algorithm uses this type of information for generating possible rules (certain and approximate). The occurrence of reducers was tested in each set. One reducer was obtained in sub-granule 4. In this reducer attribute k11 has not appeared – mathematical competence. Due to the lack of reducers in individual sets the exhaustive sets of rules in the whole study were not generated.

The next step of the research procedure involved the formulation of the minimal sets of decision rules for each granule. Table 1 contains the number of rules generated for two adopted values of the cover parameter – 0,9 and 0,5. This factor defines the expected degree of coverage of the training set by derived rules. Results obtained for the main granule and individual sub-granules, with the coverage coefficient equal 0,9, indicated the greater number of rules than when it was the 0,5 level.

The more numerous sets had rules with a smaller number of conditional attributes (min 8). However, their individual coverage were small (about 2, 5% for the most reduced rules). In both sets the degree of rule reduction with the highest coverage has not been significantly different.

Table1. Summary of test results

Scope of re- search	Number of offers	Division of the offer number – English (yes/no)	Number of the minimal rule set			
			j. covers = 0,9		j. covers = 0,5	
			possible	certain	possible	certain
Main granule	4636	1930/2706	359	237	54	40
Sub-granule 1	747	423/324	160	115	47	34
Sub-granule 2	1445	614/831	218	152	54	43
Sub-granule 3	1554	622/932	214	143	41	30
Sub-granule 4	882	268/614	114	77	21	12

Source: own study

The support rate was calculated, which is the percentage expressed indicator of the number of the set of cases confirming the rule and the indicator of likelihood (accuracy, absolute support), expressing the percentage of the number of the non-empty sub-set of cases representing the given concept (“yes” or “no”) for the decision attribute. The highest values of the support rate were as follows: in the main granule – 4,1%; in sub-granules 1, 2, 3, 4 respectively: 5%, 3,3%, 5,4% and 7,6%. In the case of the likelihood ratio we obtained: in the main granule –7,6%; in sub-granules 1, 2, 3, 4 appropriately: 6,4%, 5,3%, 8,3% and 7,8%.

SUMMARY OF THE RESULTS

When analysing the studies involving both all offers directed to specialists, coming from the whole Poland (the main granule) and the offers obtained from regions grouped in particular sub-granules, the strongest regularities deserve attention.

In the offers grouped in the main granule, if the offer of work concerned specialists for the information and communication technology apart from the computer and professional competences what was expected was the knowledge of the English language. This rule had the strongest coverage also in sub-granule 1 and 3. In other sub-granules this trend had a much lower support. While if the offers (the main granule) were directed to the specialists in economy and management, with individual and psychological competences, or when no competence was required, the knowledge of the English language was not necessary. In offers derived from areas focused in sub-granules 2, 3, 4 the knowledge of the English language was also not expected. Sometimes additional competences were required from candidates for the employee, e.g., computer, disposition or managerial. Only in the Masovia region

(sub-granule 1) it was expected that the economists, with competences: computer, individual, interpersonal and disposition, will be able to use English. Offers directed to specialists in teaching and education (living in Poland – the main granule) did not include expectations concerning skills of the English use. This language was also not required if the demand was reported for the individual, physical, interpersonal and managerial competence. This trend was confirmed by certain rules obtained from the sub-granule 2, 3 and 4. The exception is the sub-granule 1 (Masovia region), where this relation has not been confirmed. In the case of health specialists, when the competence requirements were not defined, also the knowledge of English was not necessary. This rule was not confirmed in sub-granule 1 (Masovia region). For specialists of physical, mathematical and technical sciences only the approximate rules were obtained. Both in the main granule and in individual sub-granules, if no competence was required from the candidates, then the knowledge of English was not expected in about 80% cases.

REFERENCES

- Becker A. (2014) Wykorzystanie informacji granularnej w analizie wymagań rynku pracy, *Prace naukowe Uniwersytetu Ekonomicznego we Wrocławiu* nr 328/2014, Taksonomia 23, Klasyfikacja i analiza danych – teoria i zastosowania, Wrocław, pp. 222-229.
- Bilans Kapitału Ludzkiego, bazy z wynikami V edycji badań BKL realizowanych w 2014r., <http://bkl.parp.gov.pl/dane> (24.08.2015).
- Borowik B., Borowik B., Laird S. (2009) Powiązania pomiędzy rozmytymi pamięciami asocjacyjnymi i rozmytą morfologią matematyczną, *Czasopismo Techniczne, Zeszyt nr 8*, pp. 139-157.
- Nowak A. (2013) Zbiory przybliżone w obszarze systemów ekspertowych, www.zsi.tech.us.edu.pl/~anowak/index.php?s=file_download&id=22 (08.08.2015).
- Grabiński T. (1992) *Metody taksonometrii*, Wyd. Akademii Ekonomicznej w Krakowie, Kraków.
- Grzymala-Busse J. W. (1992) LERS – a system for learning from examples based on rough sets, [w:] R. Słowiński (red.), *Intelligent Decision Support, Handbook of Applications and Advances of the Rough Sets Theory*, Kluwer Academic Publishers, Dordrecht, , pp. 3-18.
- Kocór M., Strzebońska A., Dawid-Sawicka M. (2015) Rynek pracy widziany oczami pracodawców, *Polska Agencja Rozwoju Przedsiębiorczości, Warszawa*, <http://bkl.parp.gov.pl/raporty-v-edycja-badan> (24.08.2015).
- Pawlak Z. (1982) Rough sets, *International Journal of Computer and Information Sciences*, 11, pp. 341-356.
- Gross domestic product, [pl] Produkt krajowy brutto– Rachunki regionalne w 2010 r. (2013), GUS, http://www.stat.gov.pl/gus/5840_3594_PLK_HTML.htm (08.08.2013).
- Skowron A. (1993) Boolean reasoning for decision rules generation, [w:] J. Komorowski., Z. Ras (red.), *Methodologies for Intelligent Systems, LNAI 689*, Springer-Verlag, Berlin, pp. 295-305.

- Yao Y. Y. (2005) Perspectives of Granular Computing, Proceedings of 2005 IEEE International Conference on Granular Computing, Vol. 1.
- Zadeh L. A. (1979) Fuzzy sets and information granularity, [w:] M. Gupta i in. (red.), Advances in Fuzzy Set Theory and Applications, North-Holland Publishing Co., Amsterdam, pp. 3-18.
- Zadeh L. A. (1997) Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic, Fuzzy Sets and Systems 90, pp. 111-127.