A Soft Computing Methodology based on Fuzzy Measures and Integrals for Ranking Workers Informing Labour Hiring Policies

Diogo Alves†, Faiyaz Doctor
School of Computer Science and Electronic Engineering
University of Essex
Colchester, Essex, U. K.
dv17302@essex.ac.uk,
fdocto@essex.ac.uk

Rahat Iqbal
Faculty of Engineering, Environment & Computing
Coventry University
Interactive Coventry Ltd
Coventry, West Midlands, U.K
rahat.iqbal@coventry.ac.uk

Ahmed J. Kattan
Ministry of Municipalities and Rural Affairs
Riyadh, Kingdom of Saudi Arabia
Akattan@momra.gov.sa

Effective policy-making and design for labour nationalization programmes requires a deep understanding of the factors impinging upon firms' decisions as regards the hiring of workers across different sectors of the economy, and crucially, how these factors interact in terms of either synergies or redundancies in the overall decision-making process. There is the need to develop a method that predictively determines the stability of employer-employee matches by ranking prospective and employed workers by combining information on firms, workers, and market or institutional variables. The objective of this paper is to present a methodology for transforming criteria in matched employer-employee data into a form expressing directly the variable importance for each match, that can then be used to estimate a fuzzy measure and corresponding Sugeno Fuzzy Integral to create an interpretable regression model that is able to predict the hiring patterns of firms given a pool of applicants. The SFI is explained and compared against three well-known benchmark regression methods in matched employer-employee data from the Kingdom of Saudi Arabia and shown to outperform them.

Results on calculating the variable importance with the Shapley Value derived from the estimated fuzzy measures for two selected jobs are also presented, within the scope of a larger intervention model which can be used to aid policy-makers in both designing policies and evaluating their outcome.

KEYWORDS

Sugeno Fuzzy Integral, Logistic Regression, Shapley Value, Interaction Index, Labour Market.

I INTRODUCTION

†Corresponding author.

The problem of determining the stability of matches between employers and employees has many practical applications and has been explored in the labour economics literature [4] [6] [13]. A case in point is the understanding of the factors impacting upon firms' hiring decisions that can inform and advise policy-makers that are required to intervene. The problem can be cast mathematically as a ranking of alternatives when hiring, and, within the framework of informed policy-making, it is crucial to recognize the factors determining the differential value of workers to firms and how these affect the rankings.

The ranking of individual workers by firms is impacted by firms' own conditions, worker features, and prevalent market/institutional conditions. This is a multi-criteria decision problem, where weights are to be attached not only to each of the aforementioned variables, but also to sets of variables taken together. As an example, a firm may be willing to attach a smaller weight to worker experience if there is a relatively small pool of applicants to draw from, than otherwise. One then says there is a *synergy* between workers experience and availability, inasmuch as both variables taken together have impact upon the value attributed to a particular worker.

Therefore, there is the need to develop a predictive method that explicitly models and *integrates* these interactions to better understand the impact of policies, especially if these are to be minimally invasive as regards firms' incentive structure, i.e, in the context of nudging policies.

Fuzzy Measure Theory [16] [17] is able to aggregate and weight criteria in this way. The theory is predicated upon the estimation of a function, known as *measure*, mapping every subset of the power set of the criteria of interest into a value, and then integrating a suitably constructed function with respect to it to predictively model an output of interest. This methodology requires as inputs observational data by individual i in the form of influence/satisfiability of a set of inputs k impacting upon a final valuation h_i of each individual observation [1]. These indices g_{ik} are denoted as *densities* in the literature, are bounded between 0 (not important) and 1 (extremely important) and are usually supplied by experts.

With the data in density form, there is then the need to aggregate all the information to model the outcome of interest as a function of the variable weighting. This step is obtained by empirically estimating a *fuzzy measure* assigning a weight to each set of variables. Finally, the aggregation step: these weights are all summed and the resulting value is produced. The summation, or *integration*, of the variables given their importance warrants the use of a *Fuzzy Integral*. Towards this end, both the Sugeno and the Choquet integral are used. While both have been explored in the literature [18], the latter has received more attention than the former. The availability of a fuzzy measure has the added benefit of interpretability: it renders possible the extraction of the contribution of each variable and the set of synergies or redundancies between them. The former notion is quantified through the *Shapley Value*, which ascertains the impact of the addition of a variable to a measure, while the latter comes to fruition through the *Interaction Index* [10]. Both are useful as complementary policy tools: there is value in knowing the weight attached to variables that can be eventually controlled by a policy-maker.

The literature on the use of Machine Learning methods to predict human behaviour is numerous and ever expanding [3][8]. However, papers reporting the use of Fuzzy Integrals in regression and classification tasks are not numerous. [1] compares the Choquet and the Sugeno integrals solutions to the problem of determining a passing grade for students that takes into account a set of criteria/reasons for students' failure as densities. [2] uses the Choquet integral on a problem of analyzing the factors underlying consumer choice of meat, given evaluations on properties and purchasing intention as densities. The paper also performs a comparison with a Multiple Linear regression for the same problem. In [7] the Choquet integral is used as a classifier to rank scientific journals where the results of the Shapley values is also presented. [19] illustrates the use of the R package *"Rfmtool"* developed by the authors, for the problem of the ranking of hotels by customers, where the Choquet integral is used for regression. All these papers assume that the existing input data is already in density form, supplied by either experts or questionnaires. However, that is often not the case for observed employer-employee data, which simply reports the observed variables pertaining to firms and respective workers over time. A heuristic method of adapting this kind of data, from a set of observables of matches to a set of criteria importance for each match, is explored in this paper. The procedure estimates a Multinomial Logistic Regression on the raw matched data, retrieves and standardizes the resulting marginal effects of variables by individual, and takes these as the densities that serve as inputs to estimate a fuzzy measure and consequent Sugeno Fuzzy Integral (SFI).

The procedure underlying this case study is tested on matched employer-employee data from the Kingdom of Saudi Arabia. The anonymized GOSI data covering firm and worker features was acquired from the Ministry of Labor and Social Development (MLSD) and, after matching workers with their respective firms, covers monthly observations of matches in 5 key employment sectors and firm size combinations over a period spanning two years. The MLSD is currently undertaking a policy of incentivizing the nationalization of the workforce, i.e, increasing the national participation rate, within the scope of the "Nitaqat" [11], as part of the Vision 2030 National Transformation Programme. A key problem facing policymakers in the MLSD is how to design a policy that balances incentives and deterrents for firms, taking into account that the incentive structure of firms is such that firms' and the Government's incentives are not aligned as regards the increase of the national participation rate. The system currently in place assigns a colour code to firms according to the nationalization rate of their workforce, with mixed results [11]. However, a finer-grained, data-driven insight into the factors determining the stability of matches as a proxy for worker value for firms is warranted, if a nudging policy with the double scope of minimal invasion and maximal effectiveness is to be undertaken. Thus, it is in this context that the SFI can be deployed as a predictive method. To the best of our knowledge, this is the first paper that uses the SFI to predictively rank workers and assign a degree of importance to firm, worker, and market variables.

The rest of the paper is organized as follows. Section II describes the datasets used and their conversion into density form using the proposed heuristic method. Section III briefly contrasts the four compared regression methods. Section IV presents the Shapley Values and Interaction Indices for two selected jobs, as well as an integrating scheme that links with a more complete analysis. Finally, Section V concludes.

II DATA

II.1 Structure of the raw data

For this case study, the original input data consisted of unmatched workers and firms of the Saudi Arabian economy, spanning the 24 months between May 2015 and May 2017. The firms' data was split into five representative sector/Firm size bins, and workers were then matched to their respective firms. Moreover, a set of market/institutional conditions was derived. 5 different partitions were considered: Large Firms in the Maintenance Shops and Workshops sector (henceforth, LMSW), Giant Firms in Transportation of Passengers (GTPO), Large Firms in Transportation of Passengers (LTPO), Large Firms in Collective and Social Services (LCSS), and Micro Firms in Health Services (MHSS). Each pooled sector/size combination table was further refined into tables tracking individual employer-employee matches by jobs. The number of contained occupations and employer-employee matches is reported in table 1:

Dataset	**LMSW**	**GTPO**	**LTPO**	**LCSS**	**MHSS**
#Jobs	101	39	76	118	134
#Matches	107876	69075	63822	111978	83650

Table 1: Number of jobs and matches within each dataset

The resulting 468 data tables, i.e, the total number of data tables by job over all 5 ensemble datasets, pool information about firm data, worker data, and market conditions, as described in table 2 below:

VARIABLE	**DESCRIPTION**
FIRM'S COLOR CODE	Label of the firm according to its nationalization level: 1 of 8 color bands.
AGE	Worker's age, in years
TENURE	Time in months in current occupation
SALARY	Worker's monthly wage in SAR
WORKER TYPE	One of twelve categories, pooling the two genders (Male/Female), nationality (Saudi/Foreigner) and educational attainment of worker (Basic, Intermediate, or Advanced)
MEAN WAGE OF SECTOR	Sectorial mean wage of the job
MEAN WAGE PROVINCE	Mean wage of the job in the province of the KSA the worker is registered in
SKILLS GAP	Difference between number of Foreigners and number of Saudis in the province performing the job
WAGE GAP	Difference between mean salary of Saudis and that of foreigners for the job
WORKER'S VALUE	Proxy variable built according to permanence within a firm's records and how specialized a job is.

Table 2: Variables contained in the constructed datasets by job

II.2 Data conversion to density form approach

The datasets pertaining to each job in each of the partitions were transformed into tables of densities, containing the normalized influence of each input for the observed worth of the worker in the worker-firm match. The following 4-step procedure was adopted:

Partition the Worker's Value proxy into three categories. The categorization distinguishes between workers employed for less than 6 months, those employed between 6 months and a year, and those employed for the entire two years, as Nitaqat performs a color code relabeling every 6 months.

Perform a multinomial logistic regression on the data, with the worker's value bin as the target variable, and all other variables in table 2 as explanatory variables.

Obtain the marginal effects for Representative Values (MER) for each match of the sample. Denote these as y_{ijk} (the impact of variable k on the probability that individual i is in value bin j)

Standardize both the resulting marginal effects and the original worker's value to a value between 0 and 1 expressing relative importance. The standardization formula used to derive densities was:

$$g_{ik} = \left| \frac{y_{ik} - y_{\min(k)}}{y_{\max(k)} - y_{\min(k)}} \right|$$

Letting h_i be the standardized worker value of individual i, the output of section II.2 by job is a matrix as shown in figure 1:

$$[G]_{ik} = \left(\begin{array}{ccc:c} g_{1ColorCode} & \mathrm{K} & g_{1WageGap} & h_1 \\ \mathrm{M} & \mathrm{O} & \mathrm{M} & \mathrm{M} \\ g_{nColorCode} & \mathrm{L} & g_{nWageGap} & h_n \end{array} \right)$$

Figure 1: Structure of the transformed data matrix containing the densities and standardized worker valuations.

III PREDICTIVE MODELLING OF WORKERS VALUE WITH FUZZY INTEGRATION

Given the set of matrices, one must map the input data variables into a set of weights illustrating how firms weight all the variables and combinations thereof in determining the value of a worker. This corresponds to the construction of a *fuzzy measure* λ (formally known as the Sugeno λ-measure) mapping each element of the power set of the 9 explanatory variables in table 2 to a weight into a value between 0 and 1, expressing a degree of importance, from 0, not relevant, to 1, extremely important. Mathematically, measures are characterized by boundary conditions (having no information sources is weighted by a 0, and taking all of them by a 1) and monotonicity (having more explanatory variables can never reduce the weight of a measure).

A fuzzy integral is a generalization of the concept of weighted mean, and aggregates the weight of the evidence into a single output [10]. Of the two fuzzy integrals, the Choquet integral fits the data best when no importance is attached to ordinal values of the data, whereas the SFI is the best suited for rankings [1]. We make use of the latter here. The SFI with respect to the estimated λ outputs an estimated worker value and can thus be used directly as a predictor.

IV EXPERIMENTAL RESULTS

IV.1 Model comparison

To assess the Sugeno Fuzzy Integral as a predictor, the resulting model fit (evaluated as the Root Mean Squared Error (RMSE)) was compared against the predictions of three

benchmark methods in the literature: Support Vector Regression (SVR) [5], Regression Trees (RT) [15], and Multiple Linear Regression (MLR) [20]. A 5-fold Cross-Validation for each job on each dataset was performed, and the respective RMSE on the test data for each fold was computed. The SFI requires that, first, a fuzzy measure is empirically estimated for each training fold set. Then, the corresponding SFI with respect to the estimated measure is determined for each instance on the remaining test fold. The benchmark methods can be applied without further considerations directly on the data from Figure 1. The set of minima (best RMSE for that fold among the four methods), means and medians of the RMSEs was computed.
Figure 2 plots the minima for each job in the LMSW data. The SFI outperforms the benchmark methods in a majority of cases:

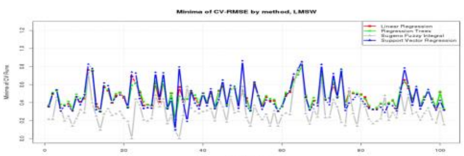

Figure 2: Minima of each cross-validated run by method for each job in the LMSW data

Table 3 summarizes the results of the comparison.

	LMSW	GTPO	LTPO	LCSS	MHSS
MLR	Median:3.96% Mean:0.99% Min:0.99%	Median:20% Mean:23% Min:12.8%	Median:15. Mean:10.5 Min:7.9%	Median:18 Mean: 14.4 Min:11%	Median: 4.5% Mean:6% Min: 6.7%
RT	Median:5.94% Mean: 5.94% Min: 1.98%	Median:12.8% Mean:12.8% Min:15.4%	Median:7.9% Mean:13% Min: 7.9%	Median:20% Mean: 21 % Min:12.7%	Median:9.7% Mean:9% Min:3%
SVR	Median:6.93% Mean:6.93% Min: 5.94%	Median:13% Mean:10% Min: 13%	Median:15.8% Mean:18% Min:12%	Median:0% Mean: 0% Min: 0%	Median:9.7% Mean:7.5% Min:7.1%
SFI	**Median:83%** **Mean:86%** **Min: 91%**	**Median:54%** **Mean:54%** **Min: 59%**	**Median:61%** **Mean: 58%** **Min: 72%**	**Median:62%** **Mean: 64%** **Min: 76%**	**Median:76%** **Mean:78%** **Min: 84%**

Table 3: Comparison of experiments of best results of each method

The cross-validation exercise outputs a set of 5 RMSE by job within each of the ensemble datasets, for each of the 4 compared regression methods. The values in each cell of Table 3 were obtained thus: First, take, for each job within each ensemble, the median, mean and minimum (1st, 2nd and 3rd rows, respectively) of the RMSE by job as metrics. Second, for each of these three metrics, the method that yielded the smallest value was determined, and the percentages obtained as the relative frequencies of best cases for that method across all jobs within an ensemble dataset. Thus, each measure sums column-wise to 1.This rough initial approach highlights the relative quality of the SFI across all metrics: the SFI had a smaller mean RMSE in 86% of all the jobs in the LMSW data, having been outperformed by the second best method (SVR), in less than 7% of all cases.

A non-parametric ANOVA (Kruskal-Wallis), coupled with a signed-rank Wilcoxon test (with the Benjamini-Hochberg p-value correction) to detect pairwise differences, all at a 5% significance level, was conducted on the full population of RMSE by method. The number of jobs for which there is a statistical difference in performance between the competing methods, and of these, for how many the SFI outperforms the others, was thus determined as those for which a statistically significant difference exists in the respectively aforementioned tests. The following table 4 reports the results:

Dataset	LMSW	GTPO	LTPO	LCSS	MHSS
#Jobs	101	39	76	118	134
#Statist. Diff	90	36	58	85	112
#(SFI > Bench.k)	76	27	49	83	97
%(SFI > Bench.k)	75%	69%	64%	70%	72%

Table 4: Statistical results on the comparison of the SFI with the other methods

The SFI outperforms the other methods in percentages ranging from 64% in the LTPO data to 75% in LMSW.

These results both match and extend those obtained in the literature. [2] reports a similar comparison against MLR and highlights the advantages of the Choquet FI over additive methods: the latter imposes no assumptions on the distribution of the error term, is less sensitive to outliers, and does not require that the feature space of the test data has the same support as that of the training data. Moreover, Fuzzy Integrals are robust to the presence of multicollinearity, i.e, the phenomenon by which there exists a linear relationship between explanatory variables. This latter issue is especially relevant in big data consisting of many different datasets with the same feature spaces. Indeed, it is likely that the variables in table 2 have different issues of multicollinearity according to the job data under consideration. All these advantages naturally carry over to the SFI.

IV.2 Shapley Value and Interaction Index

Policy insight requires a gauge on the worth of information sources, understood here as the explanatory variables affecting the hiring pattern of firms. In particular, special interest is to be attached to variables that can eventually be affected by policies, i.e, those that are not idiosyncratic to firms or workers (namely, the mean wages of the sector and province by occupation, as well as local relative skills and wage gaps), both in isolation and in interaction with specific firm and worker variables. Policy can then be fine - tuned in a way that takes interactions of variables into account.
The weight attached to each variable can be determined as the difference in the measure with and without it. The Shapley Value of a variable offers such a tool. The weight of sets of variables taken together is determined by the Interaction

Index, measuring the synergies or redundancies existing between sets of variables. A useful rule of thumb is to include in the analysis variables with a high Shapley Value and exclude those with redundancies.

As an example of application, we compute the Shapley values and two-way Interaction Indices: the Interaction index of every possible pair of variables in the estimated measure, for two chosen jobs ("Plumber" in the LMSW dataset and "General Mechanical Engineer" in the LTPO data), as shown in Figure 3 below:

	Tenure	Salary	Worker Type	Firm Type	Mean Wage Sector	Mean Wage Prov.	Skills Gap	Wage Gap	Shapley Value
Age	-0.036	0	0.0033	-0.0307	0.1474	-0.1266	-0.1766	0.2193	0.1122
Tenure		0	0.1439	0	-0.036	-0.036	-0.036	0	0.0240
Salary			0	0	0	0	0	0	0
Worker Type				-0.081	0.0273	0.0033	-0.2967	-0.081	0.2016
Firm Type					-0.0667	0.2193	-0.08	0.0394	0.0522
Mean Wage Sector						0.1474	-0.1526	-0.067	0.1002
Mean Wage Prov.							-0.1766	-0.031	0.1122
Skills Gap								-0.081	0.3455
Wage Gap									0.0522

	Tenure	Salary	Worker Type	Firm Type	Mean Wage Sector	Mean Wage Prov.	Skills Gap	Wage Gap	Shapley Value
Age	-0.076	-0.046	-0.0033	0.0121	0.0142	0.0508	0.0306	-0.0411	0.0534
Tenure		0.028	0.0093	0.0020	-0.0178	0.0999	0.0654	0.0379	0.0787
Salary			-0.0368	0.0073	0.0007	-0.0012	0.0301	-0.0396	0.0615
Worker Type				-0.1183	-0.0515	-0.0673	0.0052	-0.1192	0.1845
Firm Type					-0.0216	-0.1462	0.0118	-0.0843	0.1102
Mean Wage Sector						0.0045	0.0185	-0.0599	0.0626
Mean Wage Prov.							0.1413	-0.0312	0.2034
Skills Gap								0.0425	0.1222
Wage Gap									0.1235

Figure 3: Bilateral interaction indices and Shapley Values for "Plumber" occupation from the LMSD data (left) and "Mechanical Engineer" from the LTPO data (right).

The salary has no bearing on the worker's value for Plumbers. This is somewhat intuitive as it is to be expected that the most valuable workers within a firm would earn a higher wage. The relative local skills gap, i.e, the difference between supplied foreign and Saudi labour, is the most important variable. It has a redundancy effect when included with other variables. Hence, analyzing the skills gap in determining how Plumbers are valued, and how it could be affected by policy to bring about a different ranking of workers, is enough in this case for policy-makers. In other words, the increase in hiring of Saudi plumbers could be brought about if policy-makers act by reducing the skills gap, by, say, incentivizing Saudi nationals to take up that profession through training programmes or some such. This is a job where purely institutional variables can be analyzed and acted upon to bring about an increase in national participation rate.

For the second case, Mechanical Engineers, the Mean wage of the province they operate in, followed by the demographic type of the worker, are the most important variables. This is to be expected, since it is a job that requires higher learning. The former variable must be featured only with Age and Tenure, as there these are the only variables with synergies between them.

We have seen how the nationalization of the workforce is a function of the applying workers by firms, and that the SFI, the construction of which is predicated on a suitable measure, is a competitive tool to assess this ranking. The effect of a policy in terms of increase in the national participation rate that aligns with the MLSD's objectives can be gauged using the SFI as a predictive method coupled with the Shapley Value in the following manner:

1. Estimate a fuzzy measure on the job data, and obtain the respective SFI as a predictive method. This method can then be deployed on new data instances.
2. Obtain the jobs for which market variables have influence as measured by both the Shapley Value and Interaction Index, and simulate the ranking of workers as determined by the SFI, as the respective market variables are varied within the context of a simulated environment. In other words, the question becomes "what are the institutional parameters that bring about an increase in the ranking of Saudi Workers by firms?"
3. Obtain the respective participation rates as a time series and compare the effectiveness of policies.

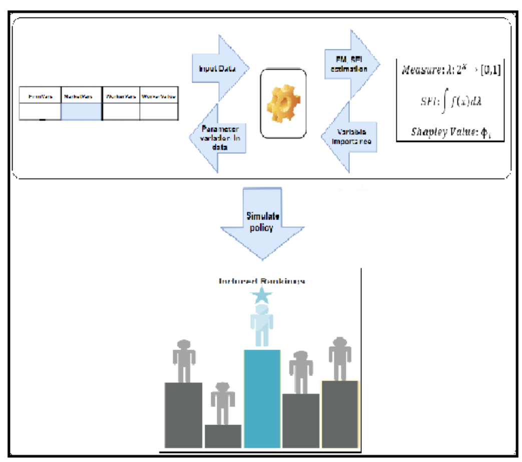

Figure 4: integration of the framework in a wider policy scope.

V. CONCLUSION

Understanding the factors that drive decision-making by firms as regards the hiring of workers is crucial from a standpoint of policy-making in the KSA's current labour nationalization programme, particularly the impact of policy-controllable variables in shaping rankings of workers by firms. Fuzzy Measure Theory is an ideal tool for this framework. However, it is not adapted to data in matched employer-employee form. This paper presents a heuristic method to convert this kind of data into tables expressing densities, using real MLSD labour data. The SFI provides a way of explaining the score of workers for firms in the KSA in a representative section of firm size/sector of the economy. The SFI is shown to outperforms the benchmark regression methods for a majority of jobs within each dataset by a wide margin.

Moreover, the construction of the fuzzy measure assigns a degree of importance to each subset of explanatory variables, which is useful as a policy-making aid, especially if the data integrates both market and idiosyncratic variables. We have explored an application in the analysis of the Shapley value and Interaction index for 2 different jobs. A sketch that integrates these into an analysis of a larger scope regarding MLSD's nationalization policy was provided.

For current and future work, we plan on tackling larger datasets by using more parsimonious methods of empirically estimating fuzzy measures from data [14][16], as well as following the procedure outlined for a complete characterization of labour market policies within the context of understanding nationalization schemes such as Nitaqat.

ACKNOWLEDGMENTS

This research was sponsored in part by a grant (Subcontract Agreement No: 203476-5103096) from the Harvard Kennedy School Evidence for Policy Design program and the Human Resources Development Fund of the Kingdom of Saudi Arabia. The research was also made possible through the support of the MLSD and access to GOSI labour market data.

REFERENCES

[1] A.M.Mane, T.D.Dongale and M.S.Bapat (2014), Application of Fuzzy Measure and Fuzzy Integral in Students Failure Decision Making, IOSR Journal of Mathematics (IOSRJM), Vol. 10, No. 6, pp. 47-53.

[2] Alfonso, L. (2013) A technical note on the use of Choquet integral to analyze consumer preferences: Application to meat consumption. J. Sens. Stud. 2013, 28, pp. 467–473.

[3] Alhabashneh, O., Iqbal, R., Doctor, F., James, A.: Fuzzy Rule Based Profiling Approach for Enterprise Information Seeking and Retrieval, Volumes 394-395, Information Sciences, 2017, pp. 18-37. doi.org/10.1016/j.ins.2016.12.040

[4] Boockmann, B. and Steffes, S. (2008), 'Workers, firms, or institutions: What determines job duration for male employees in Germany?', ZEW DP 2008-116, Mannheim

[5] Drucker, H., Burges, C.J.C., Kaufman, L., Smola, A. and Vapnik (1997), V. Support vector regression machines. Advances in Neural Information Processing Systems, 9, pp. 155–161

[6] Farber, Henry S. (1999). "The Dynamics of Job Changes in Labor Markets." Chapter 37 in Handbook of Labor Economics, edition 1, volume 3, eds. O. Ashenfelter and D. Card. Amsterdam: Elsevier, pp. 2439–2483.

[7] G. Beliakov, S. James (2011), Citation-based journal ranks: the use of fuzzy measures, Fuzzy Sets Syst. 167 (1), pp. 101–119.

[8] Iqbal, R., Maniak, T., Doctor, F., Karyotis, C.: Fault Detection and Isolation in Industrial Processes Using Deep Learning Approaches, IEEE Transaction on Industrial Informatics, 2019 (in press).

[9] M. Ayub. (2009). Choquet and Sugeno Integrals (Unpublished Master's thesis). Blekinge Institute of Technology.

[10] M. Grabisch and C. Labreuche (2004), "Fuzzy measures and integrals in MCDA," in Multiple Criteria Decision Analysis, J. Figueira, S. Greco, and M. Ehrgott, Eds. Kluwer Academic Publishers, pp.

[11] Peck, J. R. (2016). Can Hiring Quotas Work? The Effect of the Nitaqat Program on the Saudi Private Sector. American Economic Journal Economic Policy

[12] Pinar, A.J., Havens, T.C., Islam, M.A., Anderson, D.T. (2017) Visualization and learning of the Choquet integral with limited training data. 2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), pp. 1–6.

[13] Pries, Michael, Rogerson, Richard, (2005). Hiring policies, labor market institutions, and labor market flows. Journal of Political Economy 113 (4), pp 811–839.

[14] T. Havens, D. Anderson, and C. Wagner (2015), "Data-informed fuzzy measures for fuzzy integration of intervals and fuzzy numbers," IEEE Trans. Fuzzy Syst., vol. 23, no. 5, pp. 1861–1875.

[15] Torgo, L., (2010). Data Mining with R: learning by case studies. CRC Press.

[16] Torra, V., & Narukawa, Y. (2006). The interpretation of fuzzy integrals and their application to fuzzy systems. International Journal of Approximate Reasoning, 41, pp. 43–58

[17] T. Murofushi and M. Sugeno, (1991) A theory of fuzzy measure representations, the Choquet integral, and null set. J. Math. Anal. Appl., vol. 159, pp. 532–549

[18] Mesiar, R., & Mesiarová, A. (2008). Fuzzy integrals—what are they? International Journal of Intelligent Systems, 23, pp. 99–212.

[19] Vu, Huy Quan, Beliakov, G and Li, G (2014) *A Choquet Integral Toolbox and Its Application in Customer Preference Analysis.* In: Data Mining Applications with R. Zhao, Y and Cen, Y, eds. Academic Press, Waltham, Massachusetts, pp.

[20] X. Yan and X. G. Su (2009), Linear Regression Analysis (World Scientific, Singapore, 2009), ISBN-13 978-981 283 410 2; A. Hoecker et al., arXiv:physics/0703039v5.