Python Library for Consumer Decision Support System with Automatic Identification of Preferences

Jarosław Wątróbski
University of Szczecin
Szczecin, Poland *jaroslaw.watrobski@usz.edu.pl*

Aleksandra Bączkiewicz
University of Szczecin
Szczecin, Poland *aleksandra.baczkiewicz@phd.usz.edu.pl*

Iga Rudawska
University of Szczecin
Szczecin, Poland *iga.rudawska@usz.edu.pl*

Abstract

The development of information systems (IS) has increased in the e-commerce field. The need for continuous improvement of decision support systems implies the integration of multiple methodologies such as expert knowledge, data mining, big data, artificial intelligence, and multi-criteria decision analysis (MCDA) methods. Artificial intelligence algorithms have proven their effectiveness as an engine for data-driven information systems. MCDA methods demonstrated usefulness in domains dealing with multiple dimensions. One of the most critical points of any MCDA procedure is criteria weighting using subjective or objective methods. However, both approaches have several limitations when there is a need to map the preferences of unavailable experts. EVO-SPOTIS library integrating a stochastic evolutionary algorithm with the MCDA method, introduced in this paper, attempts to address this problem. In this approach, the Differential Evolution (DE) algorithm is used to identify decision-makers' preferences based on datasets evaluated by experts in the past. The Stable Preference Ordering Towards Ideal Solution (SPOTIS) method is used to compute the DE objective function's values and perform the final evaluation of alternatives using the identified weights. Results confirm the high potential of the library for identification preferences and modeling customer behavior.

Keywords: Decision Support Systems, Differential Evolution Algorithm, Preference Identification, MCDA, SPOTIS

1. Introduction

Information systems are increasingly popular in many areas. One of them is consumer decision support on e-commerce websites. Nowadays, recommender systems must consider the high accuracy of recommendations, the diversified needs of users, and the variety and novelty of products [11]. In such a situation, the problem of contradictions in the considered recommendation criteria arises, which suggests the use of Multi-Criteria Decision Analysis Methods (MCDA) to obtain a reliable recommendation [13]. However, a requirement for interactive involvement of system analysts and domain experts in the decision-making process is a significant limitation of MCDA in Decision Support Systems (DSS). On the other hand, methods for extracting knowledge from data can provide many opportunities and the essential information to improve DSS functionality.

This paper is an attempt at MCDA process automatization in the mentioned aspect. The authors' purpose was to develop a Python library for obtaining product recommendations, which could provide an engine for a decision support system recommending the advantageous product to buy. To ensure system autonomy, the authors combined the stochastic DE evolutionary algorithm with the SPOTIS method, which, based on historical datasets, identifies decision-makers' preferences represented by criteria weights. In particular, this paper introduces the EVO-SPOTIS method combining an evolutionary algorithm with the MCDA method. Such an approach makes it possible to consider many different technical attributes of products in the recommendation process incorporating identified preferences related to considered criteria. The presented approach employs a stochastic evolutionary algorithm called Differential Evolution (DE) integrated with the Stable Preference Ordering Towards Ideal Solution method (SPOTIS). The most important module of the implemented library includes the DE algorithm used to identify decision-makers' preferences. At the same time, the SPOTIS method supplies the fitness function of this algorithm. The value of the fitness function is the correlation coefficient of the real ranking and the predicted ranking, which is generated by the SPOTIS method using actual weights determined by DE. A practical application of EVO-SPOTIS is presented in the example of mobile phones.

Employing the SPOTIS method is justified by its several advantages, which include resistance to the phenomenon of reversal of rankings and the ability to identify a full domain model by establishing bounds by the decision-maker, which is essential for a wide range of product attributes. Another advantage of this method is the uncomplicated and fast algorithm, which saves time when running many iterations required by a stochastic algorithm. Instead, the application of the DE algorithm is justified by the fact that the algorithm is multi-objective, effective in continuous-based optimization problems [16], and the problem of identifying weights that are real numbers is in the continuous domain. Due to the rapid reaching of convergence by the DE algorithm, a small number of iterations is required, which saves time when multiple simulations are necessary. Furthermore, the small number of setting parameters needed and the simple algorithm with high accuracy and robustness make DE suitable for implementation to solve real-life problems for both research and practical purposes. The authors implemented the library in Python 3 and provided it in the repository of software for the Python programming language, the Python Package Index (PyPI) `https://pypi.org/project/evo-spotis/`. It can be easily downloaded and installed using pip. Complete codes and examples of usage with datasets are also available on GitHub at `https://github.com/energyinpython/EVO-SPOTIS`. The implementation includes the DE algorithm adapted to the presented problem, the SPOTIS method, and other supporting methods and solution visualization techniques. In addition, the delivered library contains complete codes of all used methods providing an environment for researchers to conduct research and simulations.

The rest of the paper is organized as follows. First, in section 2 examples of applications of stochastic algorithms and MCDA methods in the field of decision support systems in reviewed research papers are discussed. Next, section 3 introduces the methodology for this research, including the DE algorithm and the SPOTIS method. Then, in section 4 results of performed research are presented and discussed. Finally, section 5 provides conclusions and drawn future works.

2. Literature Review

Recommendation systems help support decisions that require consideration of multiple selection criteria and alternatives. These tools estimate decision-makers' preferences based on historical data, interactions, or feedback. In recommendation problems, artificial intelligence methods are commonly used, including machine learning, especially deep learning using neural networks, which can grasp nonlinear relationships in data [8, 23]. Another group of methods that can be useful in recommender systems is represented by stochastic optimization methods that search for optimal preference values of criteria by which the alternatives have been previously scored. The Differential Evolution algorithm (DE), which belongs to the group of evolutionary algorithms, is a stochastic, population-based best solution search technique. This algorithm was developed by Storn and Price [18] and is inspired by the biological model of evolution and natural selection. This technique is fast and simple, and its practical utility has been proved in continuous problems solved in many fields, such as signal processing [4], multiobjective optimization [10] and pattern recognition [3]. Among the main advantages of DE are fast convergence, a small number of set parameters required, robustness, and ease of implementation [10]. The main principle of DE is based on the iterative generation of candidate populations representing solutions using evolutionary operations such as mutation, crossover, and selection so that the best individuals can survive in subsequent generations. This procedure aims to obtain the best possible individuals whose quality is measured using a fitness function appropriately defined for the problem [1].

The potential of the DE algorithm has been successfully exploited in research focused on multi-criteria evaluation problems where preference identification is an important point. Among the examples of hybrid approaches integrating multi-object DE algorithm with MCDA methods, the work involving optimization of photovoltaic system configuration parameters is worth mentioning by minimizing technical and cost objective functions [16]. The authors of this work used the DE algorithm integrated with AHP-TOPSIS. The authors of another paper employed the DE algorithm to determine the weights of meteorological disaster evaluation indicators, and the evaluation model is based on the VIKOR method [22]. In another work, an application of differential evolution for determining meteorological disasters criteria importance is presented and discussed. Multi-criteria meteorological disaster risk assessment is based on the PROMETHEE method [21]. There is also work demonstrating the usage of the DE algorithm to generate criteria weights necessary for the Electre III method, which was aimed to produce rankings based on collected scores in multi-criteria enterprise inventory assessment [5]. Paper [14] presents a hybrid combination of the TOPSIS and DE algorithm for establishing weights in scenario evaluation in rail transportation.

Stochastic algorithms are applied to problems where the critical point is the lack of specific information necessary to solve a multi-criteria problem. The authors of paper [15] demonstrated the use of a genetic algorithm, tabu search, and simulated annealing, among others, to reconstruct incomplete rankings. In the described problem, the authors knew the incomplete preferences of dozens of decision-makers in a case study on ranking insulating materials. In the next paper [19] the application of a genetic algorithm to solve the optimization problem of online shopping is presented. The authors of this paper demonstrated the effectiveness of this approach in achieving optimal solutions for the problem in which the product's price is not the only determinant for the recommendation. A personalized recommendation system considering multiple conflicting metrics as selection criteria applying a multiobjective personalized recommendation algorithm using extreme point guided evolutionary computation (MOEA-EPG) is the topic of paper [11].

The cited examples confirm the usefulness and effectiveness of the DE algorithm in hybrid combination with various MCDA methods, focusing on identifying evaluation criteria weights. The above literature review proves stochastic algorithms' high popularity and effectiveness to support autonomic recommender systems considering multiple selection criteria. Furthermore, authors using approaches involving stochastic algorithms have noted its utility in the e-commerce domain. Stochastic algorithms particularly frequently used for this purpose include evolutionary and genetic algorithms.

3. Methodology

This paper aims to demonstrate the applicability of the methodology integrating the DE algorithm and MCDA procedure in decision support systems to identify decision-makers' preferences in a given multi-criteria problem. Thus, criteria weights are sought based on a sufficiently large historical training dataset containing performance data for analogous alternatives evaluated previously by decision-makers concerning considered criteria. The placement and practical application of the implemented EVO-SPOTIS library as the algorithmic layer of a decision support system for product recommendation is displayed as a flowchart of the system in Figure 1. The developed EVO-SPOTIS method is an integral part of the backend layer in the exemplary e-commerce system shown in the presented flowchart.

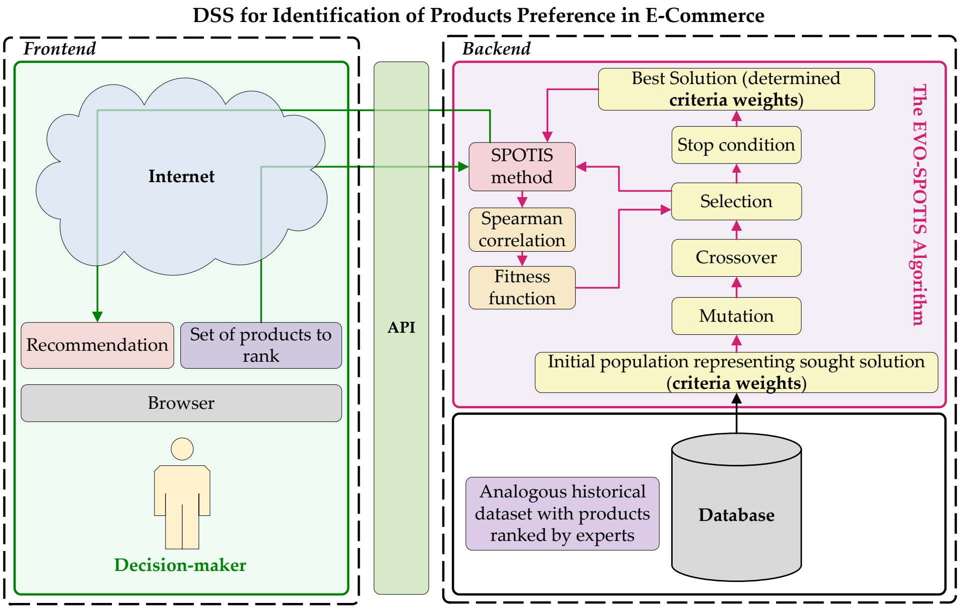

Fig. 1. Flowchart presenting the framework of DSS based on the EVO-SPOTIS library.

A training dataset for solving such a problem consists of samples representing alternatives. These samples contain performance values for considered evaluation criteria. Criteria performances play the role of training features. The target variable is defined by the real ranking determined by decision-makers for the training dataset of assessed alternatives. The test dataset is approximately 20% of the size of the whole dataset and contains performance values of alternatives for considered criteria, while the criteria weights and alternatives' ranking are unknown. Therefore, the goal is to identify criteria weights related to the training dataset using the DE algorithm and the selected MCDA method. The determined weights are then applied to generate a recommendation ranking using the MCDA method for the test dataset.

First, an initial population of a specified size is generated. Then, each individual, including N variables representing the solution under search, which are the criteria weights, evolves in successive iterations of the algorithm. Next, candidate solutions are evaluated using the fitness function. This research employs a fitness function to determine the convergence of real ranking and ranking generated by the SPOTIS method using criteria weights determined by the DE algorithm. Consistency between both rankings is measured with Spearman's rank correlation coefficient r_s. It is a profit function because the high convergence between compared rankings confirms the goodness of the solution.

3.1. The Differential Evolution Algorithm

The optimization of the problem is performed in the DE algorithm in the iterative procedure of improving candidate solutions by applying three evolutionary operators: mutation, crossover, and selection. The subsequent phases of the DE algorithm are presented in the following steps, based on [1, 18].

Step 1. Generate an initial population of a given, constant size PS. Then, each individual $x_{i,g}$ containing N variables $(j = 1, 2, \ldots, N)$ representing the solution sought, namely criteria weights, evolves in subsequent steps involving three evolutionary operators: mutation, crossover, and selection. The symbol i denotes the number of individuals, and g means the generation number.

Step 2. Mutation operation. Generate a mutated vector v_i for each vector x_i, according to the following Equation (1) according to basic strategy DE/rand/1

$$v_{i,g+1} = x_{r1,g} + F \cdot (x_{r2,g} - x_{r3,g}), r_1 \neq r_2 \neq r_3 \tag{1}$$

where $r_1, r_2, r_3 \in \{1, 2, \ldots, PS\}$ are randomly selected indexes different from the actual index i. $F \in [0, 2]$ is a mutation parameter which provides the amplification of the differential variation $(x_{r2,g} - x_{r3,g})$. The mutation can also be performed employing the DE/best/1/ strategy, as shown in Equation (2), which can increase the convergence of the algorithm using the best information found in the evolutionary process [22].

$$v_{i,g+1} = x_{best,g} + F \cdot (x_{r1,g} - x_{r2,g}), r_1 \neq r_2 \tag{2}$$

where $x_{best,g}$ represents the best individual vector in generation g.

Step 3. Crossover operation. Generate u trial vectors in the procedure of combination target and mutated vectors, namely offspring. This procedure is performed to increase the diversity of vectors. Crossover may be binomial or exponential [7]. Binomial crossover operation performed in this research involves the generation of the sample vectors according to Equation (3)

$$u_{i,g+1} = \begin{cases} v_{i,g+1}, & rand(j) \leq CR \; or \; j = a_j \\ x_{i,g}, & otherwise \end{cases} \tag{3}$$

where $CR \in [0, 1]$ is the constant crossover probability, $rand(j) \in [0, 1]$ represents the jth evaluation of a uniform random number and a_j, $(j = 1, 2, \ldots, N)$ denotes a randomly chosen index which provides that $u_{i,g+1}$ gets at least one element from mutated vectors $v_{i,g+1}$.

Step 4. Selection operation. Compare each trial vector with the corresponding target vector based on the fitness function value f. This procedure determines which individuals are expected to survive to the next generation $(g+1)$. The fitness function includes determining the ranking of training alternatives by the MCDA method using weights determined by the DE algorithm. The authors chose the SPOTIS method for the aim of this research. The vector of weights determined by the DE algorithm is used in MCDA evaluation. Then, the convergence of real ranking and ranking generated using weights determined with DE is measured by the r_s coefficient. Suppose the fitness function f value for the trial solution is better than for the corresponding individual in the population. In that case, the individual is replaced with the trial solution, as shown in Equation (4).

$$x_{i,g+1} = \begin{cases} u_{i,g+1}, & f(u_{i,g+1}) \geq f(x_{i,g}) \quad for \; profit \; f \\ u_{i,g+1}, & f(u_{i,g+1}) \leq f(x_{i,g}) \quad for \; cost \; f \\ x_{i,g} & otherwise \end{cases} \tag{4}$$

The number of DE iterations is defined or dependent on reaching a predefined fitness function value. The final result of DE is the solution represented by the best individual at the termination of the algorithm.

3.2. The SPOTIS Method

The following steps of the SPOTIS (Stable Preference Ordering Towards Ideal Solution) method are presented based on [6].

Step 1. Define the MCDA problem with the specification of minimum and maximum bounds of performance values for each criterion contained in decision matrix $S = [s_{ij}]_{m\times n}$. For each criterion $C_j (j = 1, 2, \dots, n)$ the minimum and maximum bounds of this criterion is determined respectively by S_j^{min} and S_j^{max}.

Step 2. Determination of Ideal Solution Point (ISP) denoted by $S^\star$ based on bounds determined in previous step. If for the criterion C_j larger score value is preferred, then the ISP for criterion C_j is $S_j^\star = S_j^{max}$. On the other hand if for the criterion C_j lower score value is preferred, then the ISP for criterion C_j is $S_j^\star = S_j^{min}$. The ideal multi-criteria best solution $S^\star$ is defined as the point of coordinates $(S_1^\star, S_j^\star, \dots, S_n^\star)$.

Step 3. Determine the normalized distances d_{ij} from ISP for each alternative A_i with Equation (5).

$$d_{ij}(A_i, s_j^\star) = \frac{|S_{ij} - S_j^\star|}{|S_j^{max} - S_j^{min}|} \tag{5}$$

Step 4. Calculation of the weighted normalized average distance according to Equation (6)

$$d(A_i, s^\star) = \sum_{j=1}^{n} w_j d_{ij}(A_i, s_j^\star) \tag{6}$$

where w_j denotes the weight of jth criterion.

Step 5. Creation of alternatives' ranking by sorting $d(A_i, s^\star)$ values in ascending order. The most preferred alternative has the lowest $d(A_i, s^\star)$ value.

3.3. The Spearman Rank Correlation Coefficient

The Spearman Rank Correlation Coefficient is calculated to compare two rankings x and y as Equation (7) demonstrates based on [17]

$$r_s = 1 - \frac{6 \cdot \sum_{i=1}^{N}(x_i - y_i)^2}{N \cdot (N^2 - 1)} \tag{7}$$

where N means size of vector x and y. In the presented paper, the obtained value measures predicted ranking accuracy regarding the real ranking.

3.4. The Practical Example of EVO-SPOTIS Library Application in Identification of Product Preferences

The developed approach combining DE algorithm with the SPOTIS method is demonstrated in a practical example of a multi-criteria evaluation of mobile phones. This problem represents a situation where the decision-maker has complete performance data on considered alternatives. Nevertheless, criteria weights are still needed to evaluate the dataset with the MCDA method. They can be determined by subjective methods that require the active participation of the decision-maker providing expertise in the problem being solved. However, objective weighting techniques can determine criteria weights based on the decision matrix. Besides, there are situations when decision-makers know the ranking of alternatives for an analogous situation evaluated by experts previously. Decision-makers may prefer to take advantage of knowledge regarding the relevance of particular criteria in the expert-rated problem. The authors present the EVO-SPOTIS method to determine the weights of criteria based on a training dataset with a known ranking estimated by experts in the past and archived. In this case, the training dataset contains performances of 1600 analogous variants concerning the same criteria as the actual problem to be solved and the ranking determined by decision-makers. Therefore, criteria weights are expected to be determined based on this historical dataset by the

DE algorithm. Meanwhile, 400 alternatives included in the actual set, which is planned to be evaluated, represent a test dataset. Training and test datasets were collected for this research from the mobile phone database available at `https://www.kaggle.com/datasets/iabhishekofficial/mobile-price-classification`.

The first step of the research was to perform simulations on a training set using the EVO-SPOTIS method to identify the best hyperparameters of the DE algorithm for which the results obtained are the most accurate. Results accuracy is determined by measuring the correlation of the ranking predicted by the EVO-SPOTIS method with the real ranking. First, iterative simulations were performed to indicate the most suitable population size PS and constant value of crossover probability CR. The number of iterations was set to 200, while the mutation parameter F was determined randomly from 0.2 to 0.8. After establishing the best DE algorithm parameter values, the next step involves using them to determine the criteria weights for the training dataset by the EVO-SPOTIS method. Then the test dataset is evaluated with the SPOTIS method using the vector of criteria weights obtained in the previous stage. This step gives a vector with the SPOTIS preference values of the alternatives, which have to be ranked in ascending order according to the SPOTIS algorithm. The ranking of the alternatives assessed from the test set is known for this investigation. The authors intentionally chose such a simulation dataset with a known ranking for the test set to have the opportunity to demonstrate the accuracy of the approach presented in the paper. Therefore, the authors can evaluate the convergence of real and predicted rankings at the final phase of the study using the EVO-SPOTIS method. The correlation value of the compared rankings can be calculated using the selected ranking correlation coefficient. In the case of this study, the authors used Spearman's rank correlation coefficient for this purpose.

4. Results

This section presents the research results involving the assessment of selected mobile phones with criteria weights determined using the EVO-SPOTIS library based on an analogous historical ranked dataset. The problem involves a multi-criteria evaluation of 400 mobile phones by the SPOTIS method, considering 21 criteria. These alternatives represent, for example, phones available at a given time in an online store or a set of phones with technical attribute values within the range chosen by the decision-maker. The training dataset consists of 1600 samples. The criteria taken into account in the evaluation are the technical parameters of the evaluated phones (C_1–C_{20}) and the price range within which the model can be bought (C_{21}). Information on the assessment criteria is provided in Table 1. The last column contains criteria types. The value of criterion type can be represented by Max, describing a profit criterion, or Min, denoting a cost criterion. Profit criteria are those with a maximization objective. On the other hand, the cost criteria aim is to minimize value.

The first step involved running simulations to determine the best hyperparameters of the DE algorithm. Simulation results, including 40 runs to determine the convergence of real ranking and ranking generated by the EVO-SPOTIS approach for a test dataset containing 200 samples measured by the r_s coefficient, are displayed as violin charts representing the distribution of this coefficient in Figure 2. In this simulation, the DE algorithm was trained on a training dataset including 800 samples. It can be observed that the r_s coefficient values for all the examined options are high, close to 1, which indicates the high convergence of the compared rankings and the high accuracy of the demonstrated approach. However, the best results were achieved for population size parameter PS equal to 60 and crossover probability parameter CR equal to 0.4. Thus, these hyperparameters were adopted for the final version of the algorithm.

In the following step of this research, the EVO-SPOTIS algorithm was applied to rank alternatives from the test dataset, including 400 samples using weights determined based on the training dataset with 1600 samples. Figure 3 presents charts displaying changes in fitness function values over subsequent algorithm iterations. The charts demonstrate the fast achieving convergence for the investigated problem. The fitness function achieves a very high convergence of the real and generated by the developed method ranking in the early iterations.

Table 1. Criteria considered in multi-criteria problem of mobile phones assessment.

C_j	Symbol	Name	Explanation	Goal
C_1	battery_power	Battery power	Total energy a battery can store in one time measured in mAh	Max
C_2	blue	Blue	Has bluetooth or not	Max
C_3	clock_speed	Clock speed	Speed at which microprocessor executes instructions	Max
C_4	dual_sim	Dual sim	Has dual sim support or not	Max
C_5	fc	Fc (Front camera)	Front Camera mega pixels	Max
C_6	four_g	Four G	Has 4G or not	Max
C_7	int_memory	Internal memory	Internal Memory in Gigabytes	Max
C_8	m_dep	Mobile depth	Mobile Depth in cm	Min
C_9	mobile_wt	Mobile weight	Weight of mobile phone	Min
C_{10}	n_cores	Number of cores	Number of cores of processor	Max
C_{11}	pc	Primary camera	Primary Camera mega pixels	Max
C_{12}	px_height	Px height	Pixel Resolution Height	Max
C_{13}	px_width	Px width	Pixel Resolution Width	Max
C_{14}	ram	RAM	Random Access Memory in Megabytes	Max
C_{15}	sc_h	Screen height	Screen Height of mobile in cm	Max
C_{16}	sc_w	Screen width	Screen Width of mobile in cm	Max
C_{17}	talk_time	Talk time	Longest time that a single battery charge will last when you are talking	Max
C_{18}	three_g	Three G	Has 3G or not	Max
C_{19}	touch_screen	Touch screen	Has touch screen or not	Max
C_{20}	wifi	Wi-fi	Has wifi or not	Max
C_{21}	price_range	Price	Range of price	Min

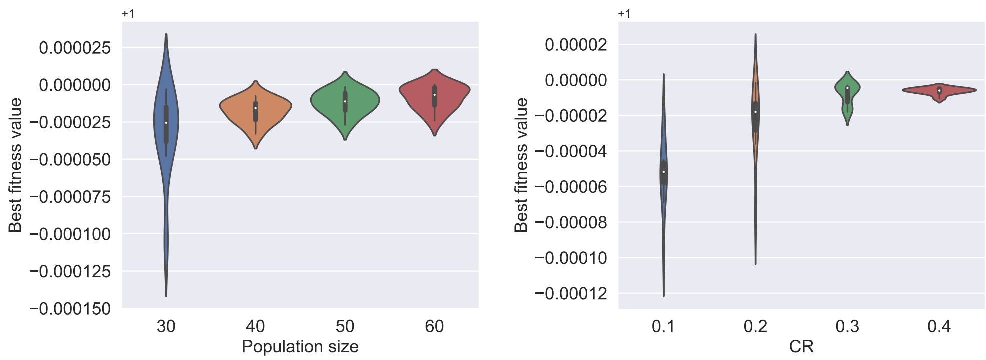

Fig. 2. Distribution of r_s values for real and predicted rankings using different DE parameters.

The comparison between the real and determined by the DE algorithm weights is visualized in Figure 4. Table 2 shows a high convergence between real and DE weights for all criteria considered. The high convergence of DE weights with the reference weights was also confirmed by the high value of Pearson's correlation coefficient, equal to 0.9999. Figure 5 demonstrates a comparison of the real ranking and the ranking generated by the EVO-SPOTIS method, considering the top 40 best-rated alternatives for a better visualization effect. A significant advantage of the results is that the top ranks are identical since the best-ranked alternatives are typically the most interesting to decision-makers. The high convergence of the predicted and reference rankings is confirmed by the high value of the r_s coefficient equal to 0.9999. It implies that compared rankings are almost identical. Complete rankings with all evaluated alternatives are available on Github.

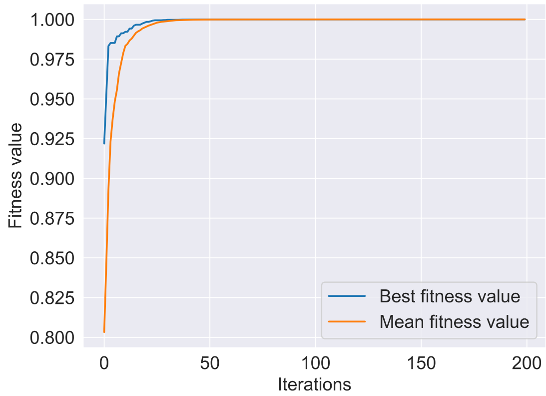

Fig. 3. Fitness function values for subsequent iterations of DE.

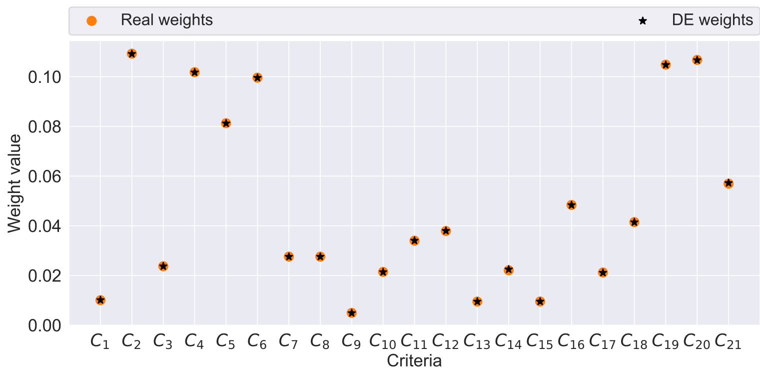

Fig. 4. Comparison of real and predicted criteria weights.

Table 2. Comparison of real weights and weights determined with DE algorithm.

C_j	Real	DE	C_j	Real	DE	C_j	Real	DE
C_1	0.0100	0.0101	C_8	0.0276	0.0276	C_{15}	0.0095	0.0095
C_2	0.1093	0.1092	C_9	0.0049	0.0049	C_{16}	0.0484	0.0483
C_3	0.0237	0.0237	C_{10}	0.0214	0.0214	C_{17}	0.0212	0.0211
C_4	0.1018	0.1018	C_{11}	0.0340	0.0340	C_{18}	0.0415	0.0414
C_5	0.0813	0.0812	C_{12}	0.0379	0.0380	C_{19}	0.1048	0.1048
C_6	0.0996	0.0996	C_{13}	0.0095	0.0095	C_{20}	0.1068	0.1067
C_7	0.0276	0.0276	C_{14}	0.0221	0.0224	C_{21}	0.0570	0.0572

The final step of the research was to compare the EVO-SPOTIS approach accuracy with other methods for objective criteria weighting. For this purpose, the authors determined rankings using the SPOTIS method for a test set of evaluated products using attribute weights determined by three different methods: Entropy [12], Criteria Importance Through Inter criteria Correlation (CRITIC) [2], and Gini coefficient-based weighting methods [20]. Then the convergence of the generated rankings with the real ranking was determined and compared, as shown in Table 3.

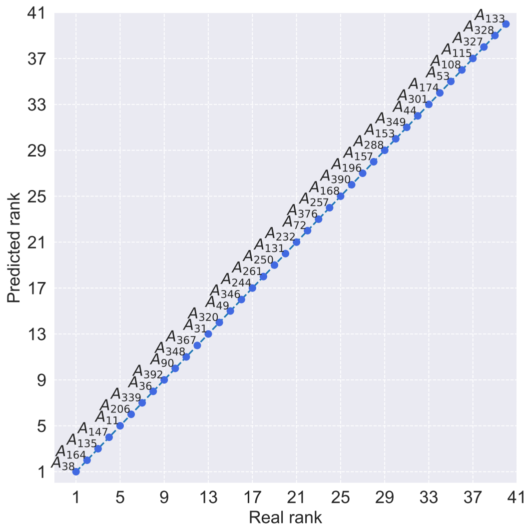

Fig. 5. Comparison of real and predicted rankings.

Table 3. r_s correlation values between rankings obtained using different criteria weights.

Weighting method	DE	Entropy	CRITIC	Gini coefficient-based
Real ranking	0.999997	0.997293	0.943665	0.952801

The results show that the EVO-SPOTIS approach enables the determination of criteria weights that are more consistent with the weights assigned by experts for the analogous historical set of products than other methods considered. The highest value of correlation coefficient r_s was achieved for the EVO-SPOTIS approach (0.9999). The obtained results confirm the reliability and high accuracy of the EVO-SPOTIS approach for identification preferences in the e-commerce domain. The high accuracy of the obtained results proves the potential usefulness of the developed library in websites of e-commerce recommendation systems concerning many attributes and technical parameters of the offered products. Furthermore, the implemented approach is an alternative to involving domain experts with each case or using objective weighting methods, especially when it is needed to personalize recommendations, for example, considering specific groups of customers. In this situation, identifying product preferences based on historical data and rankings makes it possible to give an accurate, personalized recommendation.

Stochastic evolutionary algorithms significantly extend the opportunities of MCDA because they can generate weights of considered attributes from data and thus substitute for the decision-maker. The developed approach presents significant practical and scientific potential in customer behavior modeling and product and service recommendation systems based on data mining and agent modeling. Thus, stochastic algorithms combined with the MCDA approach gain new application areas [9]. The obtained results confirm the high potential of the proposed approach in multi-criteria problems where identification of decision-makers' preferences is essential, for example, in modeling customer behavior and designing recommender systems in e-commerce. The developed library may be applicable in e-commerce as an engine for a product recommendation system to support purchasing decisions. Additionally, the library enables the induction of behavioral models of decision-makers and the determination of selection criteria guiding buyers through the identification of preferences. Furthermore, by determining the weights of the criteria autonomously, the presented library breaks the classical DSS paradigm, requiring the determination of the importance of evaluation criteria by an expert. Thus, it is an attractive and beneficial alternative to classical DSS, requiring objective or subjective criteria weighting.

5. Conclusion

The purpose of this paper was to present the usefulness and effectiveness of the developed EVO-SPOTIS method-based Python package implemented for identifying criteria weights from historical data and multi-criteria evaluation of alternatives using established weights. The presented method has the potential to be an engine for data-driven decision support systems in the field of e-commerce. The practical application of the implemented library is presented on the problem of multi-criteria assessment of mobile phones. The obtained results confirmed that the presented method gives accurate and reproducible outcomes for the investigated problem. Furthermore, the proposed approach accurately reflects the decision maker's preferences and quickly reaches convergence, so it does not require many iterations, implying an advantageous execution time. Therefore, the developed package shows high potential in autonomous data-driven recommender systems and decision support systems. Future work directions include investigations for larger datasets and exploring the effect of training dataset size on the efficiency of the proposed approach. Among future works are also investigations performed for datasets representing other domains. Future research will also explore other stochastic algorithms and their potential in the decision-makers' preference identification, for example, other evolutionary, genetic, and swarm algorithms.

Acknowledgements

The project is financed within the framework of the program of the Minister of Science and Higher Education under the name "Regional Excellence Initiative" for the years 2019-2022, project number 001/RID/2018/19, the amount of financing: PLN 10,684,000.00 (J.W., A.B. and I.R.).

References

1. Ali, I.M., Essam, D., Kasmarik, K.: A novel design of differential evolution for solving discrete traveling salesman problems. Swarm and Evolutionary Computation. 52, pp. 100607 (2020)
2. Bączkiewicz, A., Kizielewicz, B., Shekhovtsov, A., Wątróbski, J., Sałabun, W.: Methodical aspects of MCDM based E-commerce recommender system. Journal of Theoretical and Applied Electronic Commerce Research 16(6), pp. 2192–2229 (2021)
3. Baig, M.Z., Aslam, N., Shum, H.P., Zhang, L.: Differential evolution algorithm as a tool for optimal feature subset selection in motor imagery EEG. Expert Systems with Applications. 90, pp. 184–195 (2017)
4. Chakravarthi, M., Chandramohan, B.: Estimation of sampling time offsets in an N-channel time-interleaved ADC network using differential evolution algorithm and correction using fractional delay filters. In: Machine Intelligence and Signal Analysis., pp. 267–278. Springer (2019)
5. Cherif, H., Ladhari, T.: Multiple criteria inventory classification approach based on differential evolution and ELECTRE III. In: International Conference on Hybrid Intelligent Systems. pp. 68–77. Springer (2016)
6. Dezert, J., Tchamova, A., Han, D., Tacnet, J.M.: The SPOTIS rank reversal free method for multi-criteria decision-making support. In: 2020 IEEE 23rd International Confer-

ence on Information Fusion (FUSION). pp. 1–8. IEEE (2020)
7. Eltaeib, T., Mahmood, A.: Differential evolution: A survey and analysis. Applied Sciences. 8(10), pp. 1945 (2018)
8. Fang, H., Guo, G., Zhang, D., Shu, Y.: Deep learning-based sequential recommender systems: Concepts, algorithms, and evaluations. In: International Conference on Web Engineering. pp. 574–577. Springer (2019)
9. Jauhar, S.K., Amin, S.H., Zolfagharinia, H.: A proposed method for third-party reverse logistics partner selection and order allocation in the cellphone industry. Computers & Industrial Engineering. 162, pp. 107719 (2021)
10. Liang, J., Xu, W., Yue, C., Yu, K., Song, H., Crisalle, O.D., Qu, B.: Multimodal multiobjective optimization with differential evolution. Swarm and evolutionary computation. 44, pp. 1028–1059 (2019)
11. Lin, Q., Wang, X., Hu, B., Ma, L., Chen, F., Li, J., Coello Coello, C.A.: Multiobjective personalized recommendation algorithm using extreme point guided evolutionary computation. Complexity. 2018 (2018)
12. Lotfi, F.H., Fallahnejad, R.: Imprecise Shannon's entropy and multi attribute decision making. Entropy. 12(1), pp. 53–62 (2010)
13. Macias-Escobar, T., Cruz-Reyes, L., Medina-Trejo, C., Gómez-Santillán, C., Rangel-Valdez, N., Fraire-Huacuja, H.: An interactive recommendation system for decision making based on the characterization of cognitive tasks. Mathematical and Computational Applications. 26(2), pp. 35 (2021)
14. Marchetti, D., Wanke, P.: Efficiency of the rail sections in Brazilian railway system, using TOPSIS and a genetic algorithm to analyse optimized scenarios. Transportation Research Part E: Logistics and Transportation Review. 135, pp. 101858 (2020)
15. Miebs, G., Kadziński, M.: Heuristic algorithms for aggregation of incomplete rankings in multiple criteria group decision making. Information Sciences. 560, pp. 107–136 (2021)
16. Muhsen, D.H., Nabil, M., Haider, H.T., Khatib, T.: A novel method for sizing of standalone photovoltaic system using multi-objective differential evolution algorithm and hybrid multi-criteria decision making methods. Energy. 174, pp. 1158–1175 (2019)
17. Sajjad, M., Sałabun, W., Faizi, S., Ismail, M., Wątróbski, J.: Statistical and analytical approach of multi-criteria group decision-making based on the correlation coefficient under intuitionistic 2-tuple fuzzy linguistic environment. Expert Systems with Applications 193, pp. 116341 (2022)
18. Storn, R., Price, K.: Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces. Journal of global optimization. 11(4), pp. 341–359 (1997)
19. Verma, S., Sinha, A., Kumar, P., Maitin, A.: Optimizing online shopping using genetic algorithm. In: 2020 3rd International Conference on Information and Computer Technologies (ICICT). pp. 271–275. IEEE (2020)
20. Wątróbski, J., Bączkiewicz, A., Ziemba, E., Sałabun, W.: Sustainable cities and communities assessment using the DARIA-TOPSIS method. Sustainable Cities and Society p. 103926 (2022)
21. Yu, X., Chen, H., Ji, Z.: Combination of probabilistic linguistic term sets and PROMETHEE to evaluate meteorological disaster risk: Case study of southeastern China. Sustainability. 11(5), pp. 1405 (2019)
22. Yu, X., Lu, Y., Cai, M.: Evaluating agro-meteorological disaster of China based on differential evolution algorithm and VIKOR. Natural Hazards. 94(2), pp. 671–687 (2018)
23. Zhang, S., Yao, L., Sun, A., Tay, Y.: Deep learning based recommender system: A survey and new perspectives. ACM Computing Surveys (CSUR). 52(1), pp. 1–38 (2019)