Granular-Based Dimension Reduction for Solar Radiation Prediction Using Adaptive Memory Programming

Abdel-Rahman Hedar*
Dept. of Computer Science in Jamoum
Umm Al-Qura Univ.
Makkah, Saudi Arabia
ahahmed@uqu.edu.sa

Alaa E. Abdel-Hakim§
Dept. of Computer Science in Jamoum
Umm Al-Qura Univ.
Makkah, Saudi Arabia
adali@uqu.edu.sa

Majid Almaraashi
Dept. of Computer Science in Jamoum
Umm Al-Qura Univ.
Makkah, Saudi Arabia
msmaraashi@uqu.edu.sa

ABSTRACT

Estimation of the solar radiation level reaching a specific zone on the surface of earth is a crucial step in the design and planning of solar energy systems. The large number of parameters affecting the estimation and prediction processes mandates dimension reduction of the input feature space. In this paper, we address this problem for a prediction system in which uncertainties play a major role. We propose an adaptive memory programming-based approach to optimize the input feature space of a solar radiation predictor. The fitness values of reducts are calculated using granular computing. The attribute reduction concept in the rough set theory is invoked and the dependency degree is used as a fitness function. The proposed methodology is evaluated using a large environmental temporal dataset collected for regions that exhibit diverse climate conditions.

Keywords

Granular computing; adaptive memory programming; dimension reduction; solar radiation prediction

1. INTRODUCTION

Solar energy is one of the main prospective sources for renewable energy [3]. Solar radiation data is a main ingredient for the design and operation of solar energy systems [8]. So, accurate prediction of solar radiation and its components at a specific location is essential. For instance, prediction of solar radiation is important to many parties like governments, enterprises and operators for making optimal strategic plans of energy generation using different energy sources. However, short-term prediction of solar energy levels is a considerable challenge. Inaccurate prediction of solar radiation levels limits the competence of solar energy with other resources. To solve this problem, several prediction and estimation models of solar radiation have been proposed in the literature, including numerical weather prediction (NWP) and artificial intelligence (AI) models, e.g. [6] and [9]). However, there is a large number of parameters, including weather and topography variables, affects underlying prediction and estimation models. Therefore, it is crucial to get a succinct set of these parameters, or features in machine learning terminology, to improve the predictor performance and to reduce the computation cost of real-time estimation or prediction.

The parameter selection for the estimation or the prediction process usually depends on intuition and experts' choices. This selection paradigm results in a relatively large number of possible input parameters, among which several are redundant or irrelevant. From another side, the large dimensionality of the input feature space makes manual selection of the most relevant features almost impossible. So, given some rich and diverse measurements, automatic dimension reduction of the input feature space becomes a must. In this paper, we propose using the tabu search attribute reduction (TSAR) [4] as a feature selection method along with a fuzzy classifier [1] for the estimation of solar radiation levels.

The rest of the paper is arranged as follows: a review of the used methodologies is presented in Section 2, followed by a discussion of the experimental setup and evaluation in Section 3. Then, the results and technical discussion are detailed in Section 4. Finally, the paper is concluded in Section 5.

2. METHODOLOGY

The proposed method modifies the TSAR method to select the best features, and then a classifier can be used based on those selected features. The work flow of the proposed method is shown in Figure 1, and their details are explained in the following subsections.

2.1 Solution Representation

Solutions are coded as a binary vector whose dimension equals the number of conditional feature attributes. A value of 1 in that vector means the corresponding feature is in-

---

*A. Hedar is also affiliated in the Dept. of Computer Science, Assiut Univ., Assiut 71526, Egypt. Email: hedar@aun.edu.eg

§A.E. Abdel-Hakim is also affiliated in the Dept. of Electrical Engineering, Assiut University, Assiut 71516, Egypt. Email: alaa.aly@eng.au.edu.eg

GECCO'16 Companion, July 20-24, 2016, Denver, CO, USA

DOI: http://dx.doi.org/10.1145/2908961.2931648

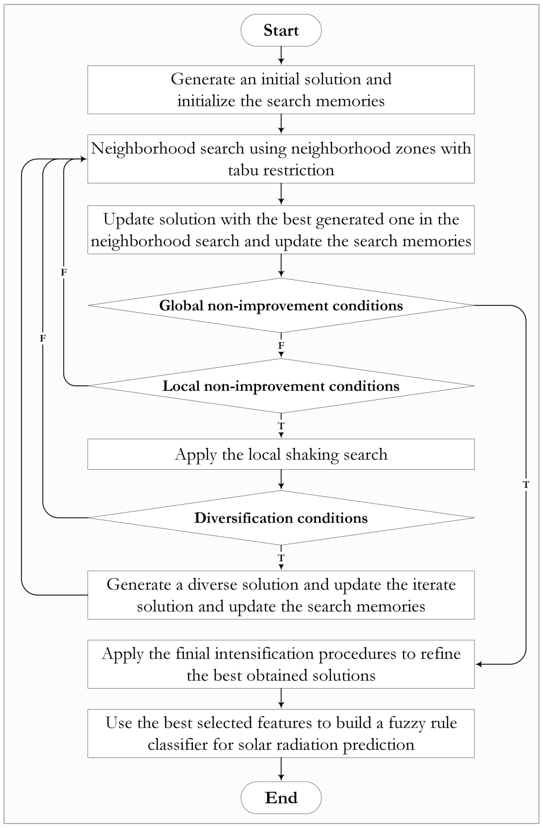

Figure 1: The flowchart of the proposed method.

cluded in that solution. Otherwise, the feature is excluded from that solution.

2.2 Solution Evaluation

In granular computing, the dependency degree function of the rough set theory [7] can be used as an evaluator to judge the solutions quality. The best solution(s) are sought through maximizing the dependency degree function and minimizing the solution cardinality. In order to compute the dependency degree function of a reduct, one can use the following definitions [7]:

DEFINITION 2.1. *For a subset of attributes $P \subseteq A$, the indiscernibility relation is defined by* $\mathrm{IND}(P)$ *[7]:*

$$\mathrm{IND}(P) = \{(\xi, \eta) \in U \times U \mid \forall a \in P, a(\xi) = a(\eta)\}.$$

It is easily shown that $IND(P)$ is an equivalence relation on the set U. The relation $IND(P), P \subseteq A$, constitutes a partition of U, which is denoted as $U/IND(P)$. If $(\xi, \eta) \in IND(P)$, then ξ and η are indiscernible by attributes from P. The equivalence classes of the P-indiscernibility relation are denoted by $[\xi]_P$.

DEFINITION 2.2. *For a subset $\Xi \subseteq U$, the P-lower approximation of Ξ can be defined as:*

$$\underline{P}\Xi = \{\xi | [\xi]_P \subseteq \Xi\}.$$

In addition, the P-upper approximation of Ξ can be defined as:

$$\overline{P}\Xi = \{\xi | [\xi]_P \cap \Xi \neq \emptyset\}.$$

DEFINITION 2.3. *The positive region of the partition of $U/\mathrm{IND}(Q)$ with respect to P, is the set of all elements of U that can be uniquely classified to blocks of the partition $U/\mathrm{IND}(Q)$ by means of P, which can be defined as:*

$$\mathrm{POS}_p(Q) = \bigcup_{\Xi \in U/\mathrm{IND}(Q)} \underline{P}\Xi.$$

DEFINITION 2.4. *The dependency degree expresses the ratio of all objects of U that can be properly classified to the blocks of the partition $U/\mathrm{IND}(Q)$ using the knowledge in P, which can be defined as:*

$$\gamma_P(Q) = \frac{|\mathrm{POS}_P(Q)|}{|U|},$$

where $|F|$ denotes the cardinality of set F.

Thus, the dependency degree is the ratio of all objects of U that can be classified to the blocks of the partition $U/IND(Q)$ using the knowledge in P.

If $\gamma_P(Q) = 1$, we say that Q depends totally on P, and if $\gamma_P(Q) < 1$, we say that Q depends partially on P. Therefore, the dependency degree $\gamma_x(D)$ of decision attribute D is used to measure the quality of a solution x. To compare two solutions x and y, x is said to be better than y if one of the following conditions holds:

- $\gamma_x(D) > \gamma_y(D)$,
- $\gamma_x(D) = \gamma_y(D)$, and the number of attributes represented in x is less than that in y.

2.3 Initialization

An initial solution is generated as a random binary vector. The tabu and elite lists are initialized as empty lists. The most recently visited solutions are stored in the tabu list in order to avoid being trapped in local optima. On the other hand, the best solutions found so far are stored in the elite list in order to be used in the intensification steps.

2.4 Search Procedures

The main search procedures of the proposed method are similar to those of our previously-published method [4] with slight modifications. Particularly, the proposed method starts with an initial solution and continues generating trial solutions. The stopping criterion is achieved when no improvement is obtained through a predefined consecutive number of iterations. Then, a diversification process begins and the search is restarted from a new diverse solution. If the number of these consecutive iterations without improvement exceeds another predefined consecutive number of iterations, an intensification process is initiated in order to improve the best reduct obtained so far. The search is terminated if the number of iterations exceeds a maximum allowed iteration limit. Finally, the search invokes a final diversification-intensification mechanism in order to obtain the final output.

2.4.1 Neighborhood Search

The neighborhood of the current iterate solution $x = (x_1, \ldots, x_n)$ is divided into a certain number of neighborhood zones Z^j, $j = 1, \ldots, \ell$. These zones are defined as follows:

$$Z^j(x) = \{x^j : x^j = (x^j_1, \ldots, x^j_n)\},$$

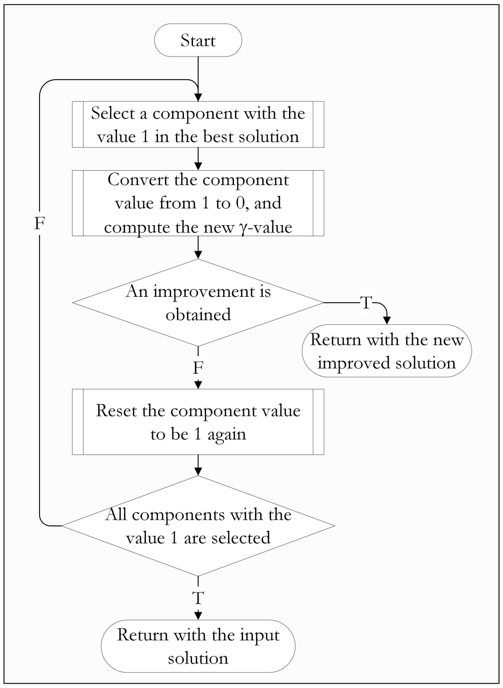

Figure 2: Reducing the cardinality of the best solution using the shaking procedure

where

$$\begin{cases} x_i^j \neq x_i, & \forall i_1, \ldots, i_j \in \{1, \ldots, n\} \text{ and } i_1 \neq \cdots \neq i_j, \\ x_i^j = x_i, & \text{otherwise.} \end{cases}$$

A trial solution is randomly generated from each zone. The tabu list is invoked to avoid regenerating recently visited solutions.

2.4.2 Solution and Memory Updates

The best trial solution among the generated ones is selected to be the next iterate. Then, the tabu and elite lists are updated.

2.4.3 Local Search

The best solution is refined using a local search procedure called Shaking [4] by trying to reduce the included attributes one-by-one without reducing its dependency degree value. Figure 2 illustrates the shaking procedure. The used shaking procedure in this paper is a modified version of the original one of [4]. The original shaking procedure [4] is only applied to reduce the cardinality of the best reducts, i.e. those whose γ-values are 1's. Here, the modified shaking procedure is called to reduce the cardinality for both total or partial reducts, i.e. for those whose γ-values equal or less than 1, respectively.

2.4.4 Diversification

Whenever diversification is needed, a new diverse solution can be generated to contain attributes chosen with probability inversely proportional to their appearance in the previously generated solutions.

2.4.5 Final Intensification

The most common features that appear in the saved solutions of the elite list are exploited as a core to generate new promising solutions. Specifically, the obtained reducts are saved in a set called Reduct Set (RedSet). The intersection of all the reducts in RedSet is called the core. Hence, a trial solution x^{Final} is constructed as the intersection of the m best reducts in RedSet, where m is a pre-specified number. If the number of attributes involved in x^{Final} is less than that in the best obtained solution by at least two, then the zero position in x^{Final}, which gives the highest γ-value, is updated to one. This update process is continued until a new better solution is found.

2.5 Control and Termination

Three non-improvement counters (I_{local}, I_{div}, I_{global}) are used to control the processes of applying the local search, diversification and final intensification, where $I_{local} < I_{div} < I_{global}$. Specifically, if a number of non-improvement iterations, I_{local}, is reached, the shaking procedure is called. Then, if the number of non-improvement iterations is increased and reaches I_{div}, a new diverse solution is generated. Finally, the final intensification is applied when the number of non-improvement iterations exceeds a limit of I_{global}.

2.6 Prediction using Classifiers

A cluster refers to a group of entities that have similar features. In fuzzy clustering, based on the theory of fuzzy sets, points can have different grades of memberships in different clusters rather than binary grades of memberships [2]. A fuzzy classifier can be constructed based on the best selected features in the previous steps.

3. EXPERIMENTAL SETUP AND EVALUATION

In order to evaluate the performance of the dimension reduction, its impact on the solar radiation estimation process is measured. We use historical observed data to estimate solar radiation. Specifically, we use weather data variables, such as temperature, humidity, wind speed and direct normal irradiance, among other environmental data, as shown in Table 2. The proposed system is evaluated by setting the objective output as the global horizontal irradiance (GHI) for the current day.

The global horizontal irradiance (GHI) measurements of historical data is used to evaluate candidate reduced feature spaces. The used datasets are collected from a number of stations in Saudi Arabia. These stations are installed and monitored by King Abdullah City for Atomic and Renewable Energy (KACARE) as a part of the Renewable Resource Monitoring and Mapping (RRMM) Program [5, 10]. The datasets represent daily observations of three Saudi cities from mid-2013 to the end of 2014. These specific datasets are used for comprehensive evaluation. However, we observed that two of the important variables were not recorded in this period due to technical issues at some of KACARE stations as they were recently installed. These two measurements are the sky cover and visibility parameters, in addition to the uncertainty values associated with all measured data. Therefore, we opted to add the visibility variable data from another source, which is the Presidency of Metrology and Environment stations' data. The choice of these cities was motivated by their different topographies and locations as well as the availability of solar data due to the research nature of the installed stations. However, only two cities of the three have the visibility parameter recorded, as shown in Table 2.

4. RESULTS AND DISCUSSION

As the main objective of this work is to optimize the input feature space, this section demonstrates the effect of this feature-space reduction on performance. Specifically, we discuss how this reduction may affect the γ-values or the prediction quality. The output space is treated in three different ways: as a continuous real-number space, as a 5-class decision space resulting from a fuzzy classifier, or as a 10-class decision space.

Figures 3-5 reflect the independent accuracy of an input attribute when each one of them is fed to the classifier as the only single attribute, representing a single-reduct input space. The three figures show the γ-values of every single attribute all three of the aforementioned output spaces. It is clear that the DH and DN attributes have the best γ-values, followed by H, while WS and PWS give the least γ-values.

The quality of dual-attribute reducts, as represented by the γ-values, is illustrated in Figures 6-8. The top-left/bottom-right diagonals of these figures are the top-view of Figures 3-5. The data non-linearity is clearly observed here. In other words, the combination of the best single attribute does not necessarily give a good reduct and vice versa. This confirms the sophistication of the considered problem.

Comprehensive result sets are shown in Tables 3 and 4 for real values and for both discrete classes, as both of the five and ten classes gave similar results. These results show how a very low-γ single-reduct attribute in combination with other attributes can provide better prediction quality. For example, combining H and DH with PWS in the KAU case raises the γ-value up to 100%. Similarity, adding the low P attribute to other attributes leads to a similar effect.

5. CONCLUSION

This paper presents an approach for feature selection and dimensionality reduction of parameters associated with solar radiation estimation. The proposed method is based on using an adaptive memory programming approach to optimize the input feature space of a solar radiation model. The correlation values of reducts are calculated using granular computing. The proposed method uses the tabu search attribute reduction (TSAR) method to select best features, and then using a fuzzy classifier based on the found features. The proposed methodology is applied to a real environmental temporal dataset collected for regions in Saudi Arabia. The results of the independent dependency degrees showed that the diffuse horizontal irradiance and direct normal irradiance attributes have the best dependency degrees values, followed by relative humidity. The dual-attribute feature reduction reveals the non-linearity behavior of the reducts, where the combination of the 'best' single attribute does not necessarily give a good reduct. Interestingly, when using discrete classes, the combination of some very low dependency degree single-reduct attributes with other attributes can lead to very good quality solutions. The results confirm the anticipated complexity of the problem, while the low number of features found highlights the need for more investigation in this area.

6. ACKNOWLEDGMENTS

This work is part of a project funded by King Abdulaziz City for Science and Technology (KACST) (grant number 13-ENES2373-10). The authors would like to thank King Abdullah City for Atomic and Renewable Energy (KACARE) for supplying the solar data.

7. REFERENCES

[1] S. Abe. *Pattern classification: neuro-fuzzy methods and their comparison*. Springer Science & Business Media, 2012.

[2] J. C. Bezdek. *Pattern recognition with fuzzy objective function algorithms*. Springer Science & Business Media, 2013.

[3] W. R. Cline. The economics of global warming. *Peterson Institute Press: All Books*, 1992.

[4] A.-R. Hedar, J. Wang, and M. Fukushima. Tabu search for attribute reduction in rough set theory. *Soft Computing*, 12(9):909–918, 2008.

[5] KACARE. Renewable resource atlas. http://rratlas.energy.gov.sa, 2015. Accessed 1st Oct 2015.

[6] T. Muneer, S. Younes, and S. Munawwar. Discourses on solar radiation modeling. *Renewable and Sustainable Energy Reviews*, 11(4):551–602, 2007.

[7] Z. Pawlak. *Rough sets: Theoretical aspects of reasoning about data*, volume 9. Springer Science & Business Media, 2012.

[8] C. Riordan, R. Hulstrom, T. Cannon, and D. Myers. Solar radiation research for photovoltaic applications. *Solar Cells*, 30(1):489–500, 1991.

[9] Z. Şen. Solar energy in progress and future research trends. *Progress in Energy and Combustion Science*, 30(4):367–416, 2004.

[10] E. Zell, S. Gasim, S. Wilcox, S. Katamoura, T. Stoffel, H. Shibli, J. Engel-Cox, and M. A. Subie. Assessment of solar radiation resources in saudi arabia. *Solar Energy*, 119:422–438, 2015.

Table 1: **The details of stations and recorded samples used in this experiment**

Station	City	Abbreviation	Latitude (N)	Longitude (E)	Elevation (m)	Data Samples
King Abdulaziz University	Jeddah	KAU	21.49604	39.24492	75	582
Qassim University	Qassim	QU	26.34668	43.76645	688	576
Taif University	Taif	TU	21.43278	40.49173	1518	575

Table 2: Solar attributes used in current experiment

Attributes	Abbreviation	KAU	QU	TU
Air Temperature (Degrees C)	T	✓	✓	✓
Average Wind Direct at 3m (Deg North)	WD	✓	✓	✓
Average Wind Speed at 3m (m/s)	WS	✓	✓	✓
Diffuse Horizontal Irradiance (Wh/m2)	DH	✓	✓	✓
Direct Normal Irradiance (Wh/m2)	DN	✓	✓	✓
Peak Wind Speed at 3m (m/s)	PWS	✓	✓	✓
Relative Humidity (Percent)	H	✓	✓	✓
Station Pressure (mB (hPa equivalent))	P	✓	✓	✓
Visibility	V	✓	✓	×

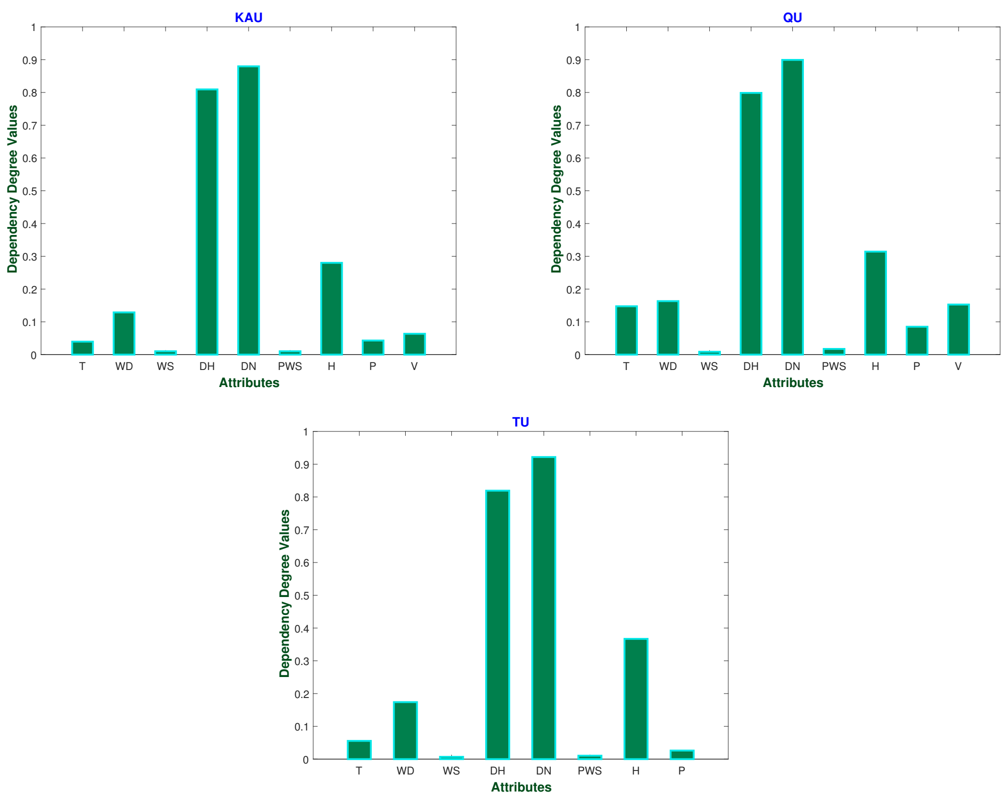

Figure 3: γ-Values of reducts with a single attribute using the real values of the decision attribute

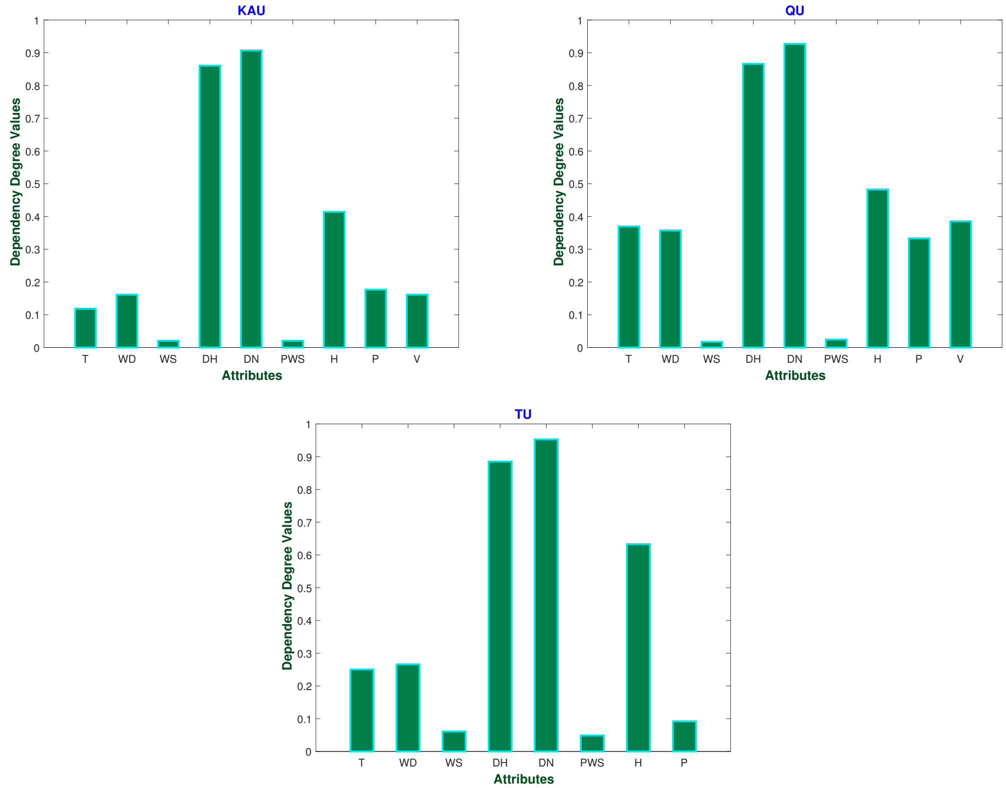

Figure 4: γ-Values of reducts with a single attribute using the 5-class decision attribute

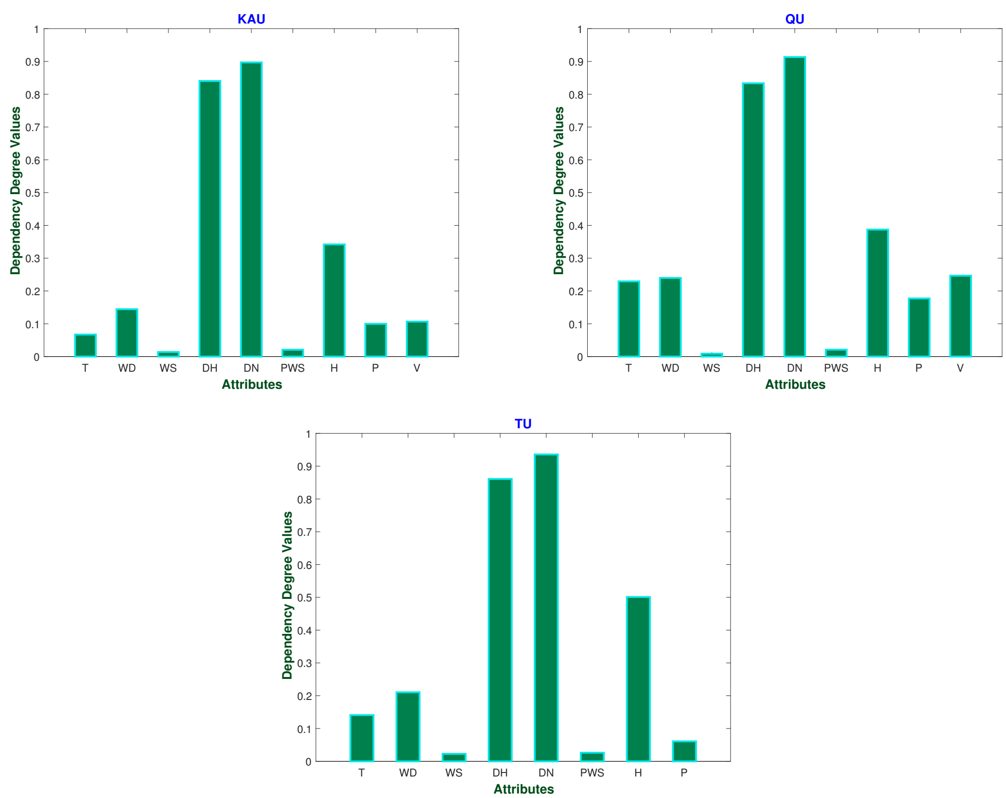

Figure 5: γ-Values of reducts with a single attribute using the 10-class decision attribute

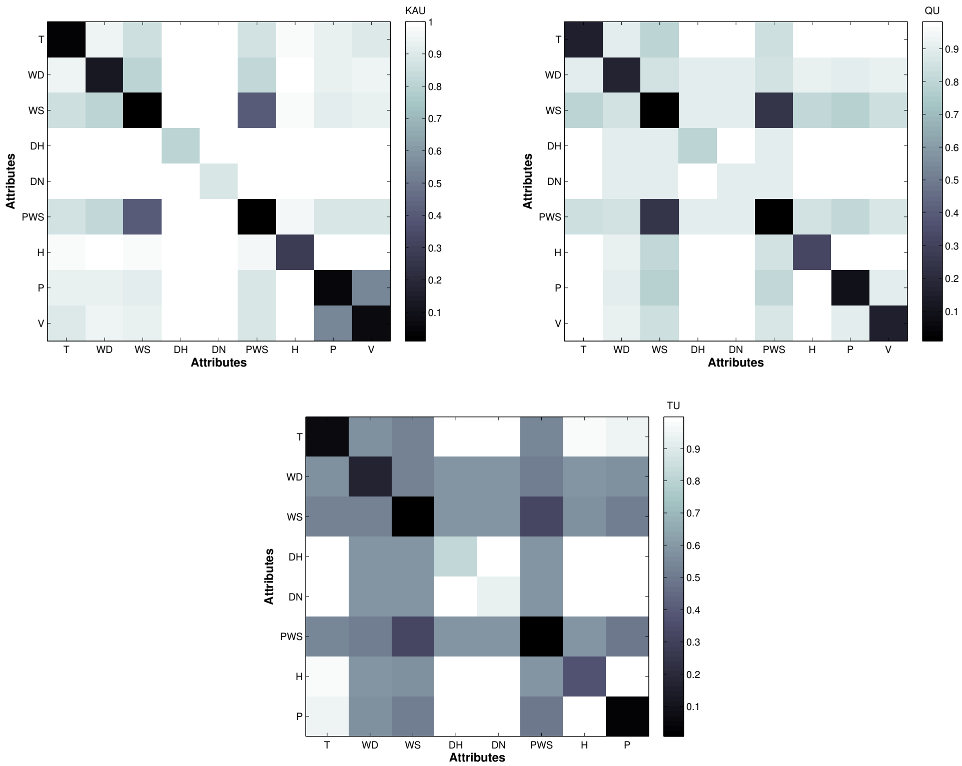

Figure 6: Distributions of γ-values of reducts with dual and single attributes using the real values of the decision attribute

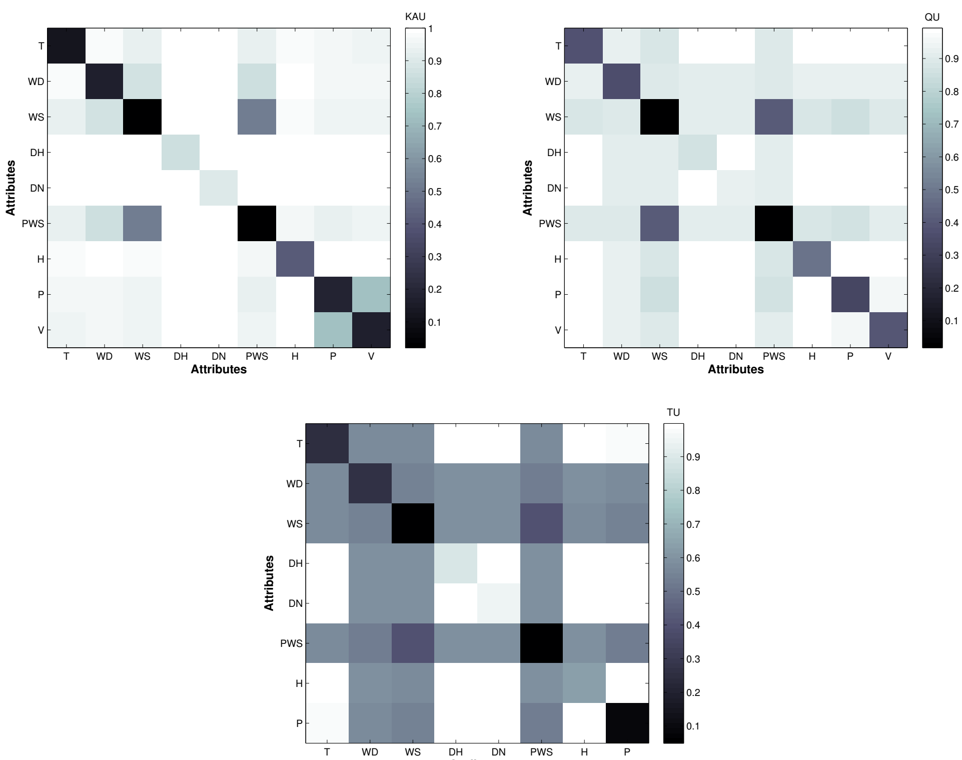

Figure 7: Distributions of γ-values of reducts with dual and single attributes using the 5-class decision attribute

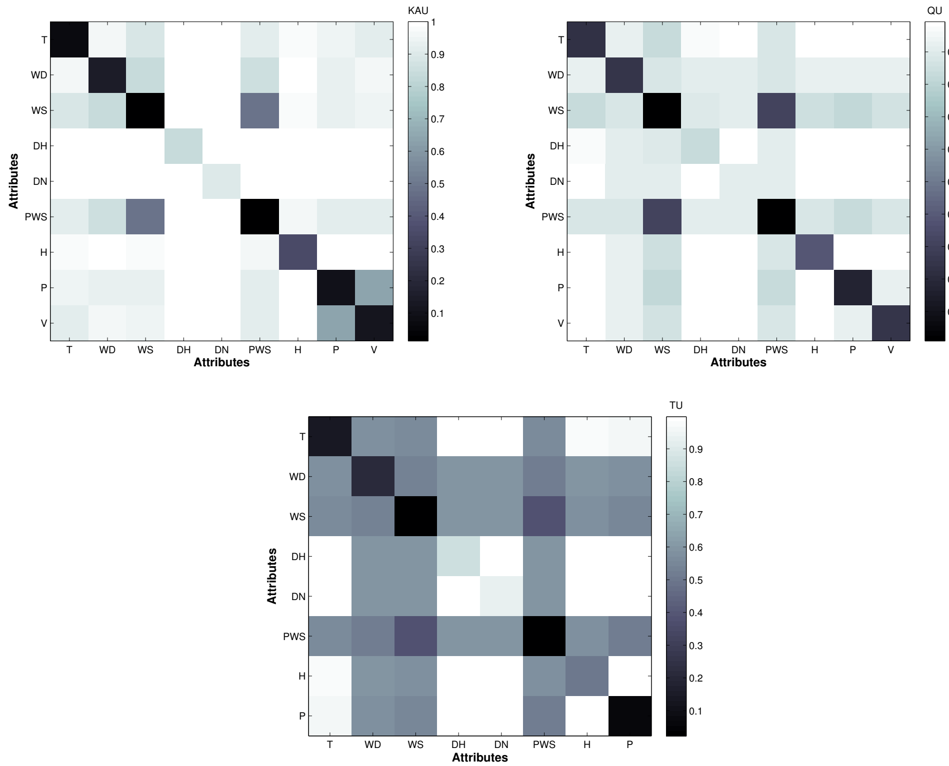

Figure 8: Distributions of γ-values of reducts with dual and single attributes using the 10-class decision attribute

Table 3: Best Reducts out of five independent runs using the real values of the decision attribute

Dataset	Attributes in the best reducts: T	WD	WS	DH	DN	PWS	H	P	V	Reduct Size	Reduct Quality
KAU				✓			✓			2	100%
QU							✓	✓	✓	3	99.65%
	✓						✓	✓		3	99.31%
					✓				✓	2	97.92%
				✓			✓			2	97.92%
TU								✓	–	1	99.83%
				✓	✓				–	2	99.83%
					✓				–	1	92.17%

Table 4: Best Reducts out of five independent runs using the 5-class or the 10-class decision attribute

Dataset	Attributes in the best reducts: T	WD	WS	DH	DN	PWS	H	P	V	Reduct Size	Reduct Quality
KAU	✓			✓						2	100%
				✓		✓	✓			3	100%
QU							✓	✓	✓	3	99.65%
							✓	✓		2	99.31%
	✓				✓		✓		✓	4	97.92%
	✓	✓					✓		✓	4	97.92%
TU								✓	–	1	99.83%