*Proceedings of the 2011 International Conference on Industrial Engineering and Operations Management*
*Kuala Lumpur, Malaysia, January 22 – 24, 2011*

# A DSS for Innovation Management: The Significance of the Innovation Determinants

**Kemal Kilic**
**Faculty of Engineering and Natural Sciences**
**Sabanci University, Istanbul, Turkey**

**Caner Hamarat**
**Faculty of Technology, Policy and Management**
**Delft University of Technology, Delft, The Netherlands**

## Abstract

In this study, a DSS for innovation management is introduced. The DSS utilizes the data acquired by a web based questionnaire which is filled in by the upper managers of the manufacturing companies. The DSS compares their answers with the responds of the other companies existing in the database of the DSS and determines the weak points of the company which have room for further improvement and at the same time significantly affect innovativeness. In order to determine the significance of the innovation determinants a genetic algorithm based feature weighting methodology that is based on k-nn classifier is developed.

## Keywords
Feature weighting, Decision Support System, Innovation Management

## 1. Introduction
This paper is part of a wider research in which a decision support system (DSS) is being developed for upper management of a company. The objective of the DSS is recommending policies in order to improve the firm level innovativeness. In the literature many factors are shown to be influencing innovation performance of companies to some extent. Those factors that influence the innovativeness are referred to as the *innovation determinants*. The DSS that's being developed utilizes an online questionnaire and assess the current status of the company in terms of particular innovation determinants. After the questionnaire is filled by the manager, the DSS benchmarks the answers of the company with the data that was collected from 184 manufacturing companies in an earlier empirical survey and automatically generates a report that recommends policies to the company management in order to become more innovative. In order to decide the scope of the recommendations that would enable the managers to focus on to the factors that have the highest marginal contribution for the company, the elimination of the least significant determinants from the analysis and identification of the significance degrees of the remaining ones is necessary.

The determination of the most significant (as well as the significance) of the innovation determinants is a onetime process and due to the relatively small size of the data the computational requirements is not the bottleneck for the problem on hand. On the other hand *feature extraction* is not an option in the very problem since we need to specify recommendations to the companies as policies in terms of certain innovation determinants. Note that, feature extraction results on reduced feature space where the new dimensions might not represent individual features any more but rather an amalgamation of various features. Furthermore, the resulting strategy recommendation will not solely base on the most significant innovation determinants but on the significant determinants that the company has also room for improvement. That is to say, if the company performs best in the most significant innovation determinant, then there is no logic of proposing a further improvement in those factors. Hence the resulting strategy recommendation should be based on the combination of two concepts, namely, *the significance of the determinants* and how the company performs in those dimensions.

In order to realize this goal, we have to reduce the number of innovation determinants available in the literature (in our study we have 32 controllable determinants) and specify the significance degrees of the remaining ones. That is to say a *hybrid* feature selection and feature weighting approach is required in the development process of such a DSS.

A simple iterative elimination process of the features based on a learning algorithm is proposed for the *feature selection* stage. The proposed learning algorithm for the selection process utilizes a *feature weighting* approach. A genetic algorithm based methodology that utilizes a *k-nn* classifier is developed for this purpose. The proposed algorithm was validated earlier [1] and utilized as the engine of the DSS.

In the next section we will introduce the architecture of the DSS. Next we will briefly review the literature on feature selection, feature weighting and the genetic algorithms which are the theoretical concepts that are relevant for the focus of this paper. In Section 4 we will present the significance degrees of the innovation determinants as the result of the analysis conducted on a real life data collected during an earlier study. We will finalize the paper with our concluding remarks and future research agenda.

## 2. The Architecture of the DSS

The proposed Decision Support System is a web-based tool that will assist the firms while they are developing policies to be more innovative. The DSS automatically generates two different reports based on the questionnaire that is filled by the company managers. First report benchmarks the company with others and the second report makes policy suggestions to the company in order to be more innovative.

The questionnaire is basically the online version of the questionnaire utilized in [2]. The benchmark reports and policy suggestion reports will initially be generated based on the existing data collected from the 184 companies. A three-layered architecture is developed for the DSS which is depicted in Figure 1. In the first layer, the user will fill out the questionnaire from his/her personal computer via a web browser (Internet Explorer, Firefox, Netscape, i.e.) and answers of the user will be sent to the second layer, which is the server.

The server will constitute the rule base (the engine). The engine will combine the significance degrees of the innovation determinants and the weakness scores of the company on the innovation scores and yield an *Overall Combined Score* (OCS). The significance degrees will be generated by the Fuzzy System Modeling Based-Feature Weighting Algorithm discussed in [1]. On the other hand the weakness scores of the particular company will be measured based on the difference between the questionnaire filled by the company manager and the responds of the companies' existing in the database, which is the third layer of the DSS.

The benchmark report will be solely based on the answers to the questionnaire and the responds of the companies' existing in the data base. It will provide various charts as line charts, radar charts and relevant information to map the company's status among the others that are in the same industry as well as all companies. On the other hand, the policy suggestion reports will be prepared by using the overall combined score (OCS) and will provide policy suggestions only on the top three innovation determinants. The Application is developed in *.Net* 3.5 Platform with C #. The Graphical User Interface developed by utilizing the ASP.NET MVC ve AJAX Technologies ASP.NET Chart Controls is used for reporting.

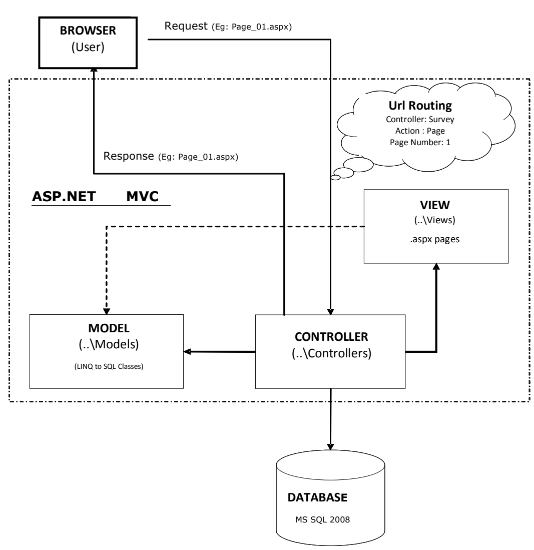


Figure 1: Architecture of the proposed Decision Support System

## 3. Relevant Literature

Note that in order to generate the strategy reports the Overall Combined Score (OCS) should be calculated for the company, thus the significance degrees of the innovation determinants should be determined. The focus of this paper is the significance degrees of the innovation determinants that are obtained from the feature weighting methodology which will be at the core of the DSS. The feature weighting problem is closely related to the feature selection (or feature subset selection) problem where there is an extensive research available. Many of the feature weighting approach roots to the existing feature subset selection methodologies so we will first cover the relevant feature subset selection literature.

### 3.1 Feature Subset Selection

In the context of classification and supervised learning problems, the *feature subset selection* is determination of the feature subset that will provide the best classification accuracy amongst the features that are required for determination of the interrelationships throughout data, namely unraveling the hidden relationships.

In the literature there are various approaches to the feature subset selection problem. These methods can be classified under three main categories, namely, the *filter methods,* the *wrappers methods* and the *embedded methods*. Briefly speaking, the *filter methods* determine the feature subsets based on an *a priori* criteria such as correlation, entropy or information acquisition directly from the sample data instead of using learning algorithms. On the other hand, the *wrappers methods* utilize various learning algorithms. The fundamental motivation of the wrapper methods is the fact that the feature subset chosen by the filtering step is the *best* subset selection in terms of the objective function of the particular filtering technique. However, generally the main objective of the feature subset selection process is to obtain the features that will provide the *best classification accuracy*. Hence, although wrappers method is more costly than the filter method in terms of the computational time, it can yield more efficient results in practice. In *embedded methods*, classification process is computed synchronously with feature subset selection.

Regardless of the method, there are some common structural elements that should be considered in feature subset selection process. These are the starting point, the searching strategy, feature subset evaluation, and the termination criterion. The starting point depends on the searching strategy. Forward insertion and backward elimination strategies are common search strategies. In the case of backward elimination one starts with a large candidate set of features and eliminates specified features (with respect to a criterion, *i.e., feature subset evaluation*). On the other hand, the forward insertion is just the opposite of backward elimination and starts with an empty subset, and at each iteration, specified features are inserted to the subset. Note that, apart from starting with an empty subset or full subset, backward elimination or forward insertion can also be applied on a subset composed of a random number of features chosen as starting point. Searching strategy is a major step in feature subset selection. Evaluation of all possible feature subsets can be extremely costly and misleading. For example, in a data set of $n$ features, computational time for searching all possible $2^n$ feature subsets would be too costly. The computational complexity of the problem has triggered the need for more efficient searching strategies.

After generating a candidate feature subset, a pre-determined subset evaluation method should be applied in order to decide which subset to use. In case of filter methods the evaluation is based on a pre-specified function. In case of wrappers method it is attained by measuring the data classification performance of the features. Finally, one should also define a termination criterion. It should be specified when or according to what criterion the applied method should be terminated. There are several options for determining the termination criterion. Terminating the algorithm when feature subset size reaches a pre-specified number or when candidate feature subsets starts to make no difference in terms of the classification accuracy are some examples of possible termination criteria.

Let's now discuss the previously mentioned three different feature subset selection approaches in more detail. Note that, due to the wide range of the literature on feature subset selection, a comprehensive literature review is not possible in a limited space. Therefore, the literature will be covered mostly conceptually and refer to only prominent research.

#### *Filter Methods*

The feature subset selection method that determines the significant features directly from the data based on a pre-specified function that indirectly measures the strength of the relationship between the features and the independent variable is referred to as the *filter methods*. Its distinctive characteristic is to filter the features directly in the feature subset generation stage and not to use the learning algorithm inside evaluation function. Learning algorithm is applied as a separate second step, after feature subsets are generated, if need be for a particular application. There are various widely used filtering methods in the literature. Most of these methods are based on statistical metrics. Among these metrics, correlation, distance between means, impurity, entropy and mutual information acquisition can be listed as examples. Apart from these statistical metrics the widely acknowledged RELIEF [3] algorithm and its improved version, namely, RELIEF-F [4] utilizes a slightly more complex process to filter the features.

Note that, as we discussed earlier by choosing different searching strategy, feature subset evaluation and termination criteria, distinct filtering methods can be obtained that utilize the same *filtering metric.* One possible method is to use forward insertion in order to add a specific number of features to the subset based on, say, their correlation with output score by adopting a nearest neighborhood approach. Another method can be picking the feature having the highest correlation with output score and eliminating specific number of features having the highest correlation with the chosen feature. Next, among the remaining features, those that have the highest correlation values with the output score might be selected by forward insertion.

***Wrapper Methods***

The wrappers methods have been introduced as an alternative for filter methods. The wrappers methods are more costly than the filter methods because feature subsets are evaluated by the learning algorithm itself. The loss of time due to the complexity is usually covered by more efficient feature selection. In wrappers method, learning algorithm is repeated for each candidate feature subset and model evaluation is performed.

Wrappers methods should answer three basic questions, namely, the strategy that will be used for searching the feature space, the performance metric that will guide the searching strategy and the classification method that will be used. In the literature, different suggestions have been proposed for each one of these questions and distinct combinations of these propositions have resulted in various methods. A detailed comparative analysis of these methods has been conducted by Kohavi and John [5].

***Embedded Methods***

Embedded methods are third type of the feature subset selection methods apart from filter methods and wrappers methods, although they are not used as common as the other two methods. In embedded methods, classification and feature subset selection is performed simultaneously. Since classification and feature subset selection is implemented simultaneously, they give efficient results in terms of time complexity. Decision trees can be given as examples of embedded methods.

### 3.2 Feature Weighting

Feature weighting aims to identify the relevance or significance degree of the features rather than eliminating them as in the case of the feature subset selection approaches. In one sense the feature subset selection problem is a special instance of the larger feature weighting problem. As opposed to specifying a feature as significant or not, i.e., assigning a 0 or 1 to the features, in the case of feature weighting a significance degree is assigned which is usually a real number from the interval [0,1]. In many data mining applications which incorporate feature weighting, the summation of the total weights of the individual features is set to be equal to 1. Such normalization of the weights provides further information regarding to the relative significance of the feature to the analyst.

Considering the fact that feature subset selection is an instance of feature weighting many of the feature subset selection algorithms discussed earlier can be used as feature weighting as well. For example while utilizing the *filter methods* rather than eliminating the features that have smaller scores in terms of the filtering functions, one can use these scores in order to identify the feature weights. This can be achieved by normalizing the scores associated with each feature and adopting them as the weights of the features either directly or after some transformations.

Similar approach is possible for the wrapper and the embedded methods. Recall that the wrapper methods determine the feature subsets based on the classification accuracies with various different strategies in terms of the search of the feature space, the performance metric to guide the search space and the classification methodology. No matter which strategies are adopted rather than elimination of the features (i.e., assigning a 0 for the features) or picking features as part of the subsets (i.e., assigning 1 for the features) weights that are real numbers between 0 and 1 can be specified at each iteration.

Feature weighting is less common in the literature due to the fact that it doesn't aim to reduce the size of the feature space. However, in our case we should specify the weights of the innovation determinants since it is going to influence the recommendations of the DSS that is being developed. Note that if the innovation determinants are only specified as relevant or irrelevant, in that case DSS would base its decision only to those determinants where there is room for further improvement. However the overall score should be a combination of both of the factors, namely the *significance of the determinant* and the *possible room for* improvement. Hence the feature weighting approach is preferred particularly in such applications

### 3.3 Genetic Algorithms

In this study we propose a Genetic Algorithm based feature weighting approach that utilizes a *k-nn* classifier. Note that the Genetic Algorithms (GA) is a global optimization tool, proposed by Holland [6] and used in many applications in the literature. Genetic Algorithms is inspired from the biological evolution process. Although there are various GA applications that differentiate from each other in many different details, basically there is a common dominant approach among all applications.

Generally speaking, the first step in GA requires constructing an initial gene pool that is composed of chromosomes, where each chromosome is a solution or representation of a solution. After the construction of the initial gene pool, the process of stepwise creation of new generations starts. At each step, next generation is obtained by using the current gene pool. The basic four methods that are utilized to create the next generation are referred to as crossover, mutation, elitist reproduction and immigration operations, which are essential part of natural biological life. Crossover is achieved by crossing pairs of chromosomes which can be conducted in various ways. On the other hand, mutation is some sort of perturbation of the chromosome in order to ensure *diversification* of the search mechanism. Furthermore some of the chromosomes from the current generation are directly transferred to the next generation. This process is referred to as the elitist reproduction. In some applications randomly created new chromosomes are introduced to the next generations, i.e., immigration process. As a result of these operations, more competitive and strong chromosomes in the gene pool are preserved for the next generations and weaker and less competitive ones are eliminated. At this point, the fundamental decision to be made is how to determine which chromosome is stronger and which one is weaker. This is done by using a *fitness function*, where each chromosome has a *score* according to this *fitness function*. This process is repeated until a termination criterion is met. The elimination process with respect to the fitness function ensures the *intensification* of the search procedure.

## 4. Analysis and Results

In this study we employed a Fuzzy System Modeling Based Feature Weighting Algorithm (FSMB-FWA) in order to infer innovativeness of the test data based on the model constructed by the training data [2]. We utilized the *k-nearest neighbor* based inference algorithm introduced in [7]. According to this approach the closest *k* neighbors are determined based on a *weighted* Euclidean distance metric. The *weights* are basically the *significance degrees* of the associated input features, i.e., the innovation determinants. The significance degrees are learned from the training data via a genetic algorithm. In the developed genetic algorithm, a chromosome structure that represents significance degrees associated with the innovation determinants is adapted. On the other hand, the average prediction error that reflects the cumulative difference between the real innovativeness scores of the training data and the inferred innovativeness scores (i.e., *prediction accuracy*) that are obtained by using the Fuzzy Inference algorithm is used as the *fitness function score* of each chromosome.

The above described technique is first tested on various benchmark data (such as *Iris Data*, *Concrete Compressive Strength Data*, etc.) which are publicly available in University of California at Irvine Machine Learning Repository [8]. The results suggested that the proposed approach correctly identifies relevant features and yields high predictive performance[1]. After the validation phase the developed FSMB-FWA is applied to the empirical data collected in [2] in order to identify the significance degrees of the determinants of innovation.

Note that, the factors that affect the innovation performance, which are referred to as the *determinants of innovations,* can be grouped into two as indigenous and exogenous factors. The indigenous factors include *general firm characteristics* such as age, size of firm, ownership status, *intellectual capital* that consists of human, social and organizational capital, *organizational culture* that includes centralization, formalization, communication, reward system, etc.; *collaborations*, *innovation outlay*, *business strategies* such as the manufacturing, marketing and technology strategies. On the other hand the exogenous factors can be referred to as the *environmental conditions* such as market dynamism, public incentives, internal and external barriers to innovations, and tax rebates. Even though these factors are demonstrated to have influence on firm level innovativeness in the literature, some of them should be eliminated directly from the list for our purpose. The factors those are not controllable by the managers (e.g., age of the company, size of the company, existence of foreign investment, etc.) were eliminated. Furthermore, we need to rate them and determine which factors are most significant due to the reasons explained earlier.

The result of the analysis conducted with the FSMB-FWA is depicted in Table 1. Note that, in the original model there were 32 determinants of innovations. In Table 1, the most significant ten innovation determinants that the managers can alter are presented. The developed DSS also checks the current status of the company in terms of these innovation determinants and compare them with other companies in order to determine if there is space for further improvement. Based on the Overall Combined Score (OCS) which is the combination of the significance degrees of the innovation determinants and weakness score, the DSS selects the top three determinants and suggest the company to focus on improving these factors.

Table 1: Significance degrees of innovation determinants

| Innovation Determinants | Significance Degrees |
|---|---|
| Organizational Capital | 0.14 |
| Monitoring Inner Milleu | 0.11 |
| Management Support | 0.11 |
| Reward Systems | 0.10 |
| Monitoring Outer Milleu | 0.10 |
| Human Capital | 0.09 |
| Social Capital | 0.09 |
| Communication | 0.09 |
| Autonomy | 0.09 |
| Monitoring Sci&Tech | 0.08 |

## 5. Conclusion

In this paper we presented a DSS framework that will assist upper managers to develop policies in order to be more innovative. The DSS relies on an online questionnaire and generates two reports based on empirical data that was collected in an earlier study [1]. One of the reports benchmarks the company with others in various innovation determinants such as firm culture (management support, reward system, etc.), intellectual capital (human capital, organizational capital, etc.), monitoring strategies, collaborations, business strategies, etc. The second report provides policy suggestion to the manager. The second report will be based on two dimensions. Firstly it should limit its attention to only those factors that the company has room further improvement and secondly the significantly improve the innovativeness of the company. This paper focuses on the second part and identifies the significance degrees of the innovation determinants. Note that this is a rare study (if not unique) that ranks and rates various innovation determinants which is essential for the engine of the DSS.

At the moment the development process of the DSS is at its final stages and the pilot studies are about to begin. After the pilot studies the DSS will be publicly available. Various different learning algorithms can be tested in order to identify a better set of significant innovation determinants and their significance degrees.


## Acknowledgements

This study was supported by a grant from The Scientific and Technological Research Council of Turkey (TUBITAK-SOBAG-107K521). We would like to thank to Prof. Gündüz Ulusoy, Prof. Lütfihak Alpkan, Prof. Cagri Bulut and Gürhan Günday for their contribution in the initial phase of this research.


## References


1. Hamarat, C., and Kilic, K., 2010, “A Genetic Algorithm Based Feature Weighting Methodology,” Computers and Industrial Engineering, Awaji Islands, Japan.
2. Gunday, G., Ulusoy, G., Kilic, K., and Alpkan, L., 2008, “Modeling Innovation: Determinants of Innovativeness and the Impact of Innovation on Firm Performance,” Proc. of the 4th IEEE International Conference on Management of Innovation and Technology, Piscataway, New Jersey, 776-771.
3. Kenji, K., and Rendell, L.A., 1992, “A Practical Approach to Feature Selection,” Proc. of the 9th International Workshop on Machine Learning, 249-256.
4. Kononenko, I., 1994, “Estimating Attributes: Analysis and Extensions of RELIEF,” Machine Learning: ECML-94, 171-197.
5. Kohavi, R., and John, G.H., 1997, “Wrappers for Feature Subset Selection”, Artificial Intelligence, 97(1-2), 273-324.
6. Holland, J.H., 1975, “Adaptation in Natural and Artificial Systems”, Univ. of Michigan Press, Ann Arbor
7. Kilic, K., Sproule, B. A., Turksen, I. B., and Naranjo, C. A., 2004, “Pharmacokinetic Applications of Fuzzy Structure Identification and Reasoning,” Information Sciences, 162, 121-137.
8. Blake, C. L., and Merz, C. J., 1998, UCI Repository of Machine Learning Databases. Univ. California, Dept. Inform. Comput. Sci., Irvine, CA. [Online]. Available: http://archive.ics.uci.edu/ml/datasets/Iris