Preface

There are many factors that motivate the involvement of a dimensionality reduction (DR) step of knowledge discovery in database (KDD) tool in a variety of problem solving systems. Many application problems process data in the form of an assemblage of real-valued vectors (for example, protein classification, bookmark categorization, text classification, etc.). If these vectors represent a high dimensionality, then processing usually becomes infeasible. So, it is sometimes useful, and often necessary, to reduce the data dimensionality to a more manageable size with a little or no information loss. Sometimes, high-dimensional complex process can be guided by considerably fewer, simple variables. The process of dimensionality reduction acts as a tool for modelling these phenomena, enhancing their clarity. There is often a considerable amount of redundant and/or irrelevant or misleading information present in the various information systems. This requires to be eliminated before any further processing can be carried out. For example, the problem of generating classification rules from large volume datasets usually get benefitted from a data reduction preprocessing step. This reduces the time required to perform induction as well as it makes the resulting rules more comprehensible and can improve the resulting classification accuracy. Semantics-destroying dimensionality reduction approaches irreversibly transform data, whereas semantics-preserving DR techniques (such as feature selection) tries to retain the meaning of the original feature set. The main objective of feature selection is to determine a minimal feature subset from a problem domain while retaining a suitably high accuracy in representing the original features. There are sometimes many features engaged to select from combinatorially enormous numbers of feature combinations. It might be anticipated that the involvement of an increasing number of features would increase the likelihood of including sufficient information to distinguish between classes. Unfortunately, this is not true if the size of the training dataset does not increase rapidly with each additional feature inserted. This is known as curse of dimensionality. A high-dimensional dataset increases the possibilities that a data-mining algorithm will discover spurious patterns that are not valid in general [22; 42]. Most of the approaches apply some degree of reduction in order to deal with large amounts of data, so a well-organized and an effective reduction method is required.

A technique that can reduce dimensionality using information contained within the data set and that preserves the original meaning of the features (i.e., semantics-preserving) is clearly desirable. Rough set theory (RST) can be applied as such a tool to find data dependencies and to reduce the number of features contained in a dataset using the data alone, requiring no additional information. Over the past ten years, RST has indeed become a topic of great interest to researchers and has been applied to many domains. The classical rough set model, introduced by Pawlak, has been effectively applied as a feature selection or attribute reduction and rule learning tool. In the rough set model, a crisp equivalence relation as well as crisp equivalence classes is applied to define the dependency function between decision and conditional attributes available in the information system. The dependency function is effectively used to establish the relevance between the decision and conditional features and to assess the classification potential of the features. However, the classical rough set model could not apply directly on the real-valued datasets due to its limited requirement of nominal data. Therefore, many generalizations of rough set model have been presented to avoid the information loss.

Fuzzy rough set [20; 21] is one of the most efficient extensions of rough set, which can be directly applied to the real-valued datasets without any modification in the information system. Predominately, fuzzy rough set can efficiently tackle both fuzziness and vagueness available in the datasets with continuous features [39; 42]. By combining rough (as proposed by Pawlak) and fuzzy sets (as proposed by Zadeh) as presented by Dubois and Prade allows the notion of fuzzy rough sets, which gives a powerful means of dealing with the problem of discretization and can be effectively implemented to the reduction of continuous attributes. In the proposed model of fuzzy rough sets, a fuzzy similarity relation is characterized by real-valued conditional features and is applied to evaluate the similarity between two objects. Fuzzy rough sets are mostly implemented to direct the inconsistency between conditional attributes and decision attributes, i.e. a few samples have similar or having the same conditional attribute values but distinct labels. With lower approximations in fuzzy rough sets, each sample can be assigned to a membership in the form of a decision class to evaluate this inconsistency, and feature selection techniques based on fuzzy rough sets focus to obtain a reduct to retain the membership of every sample. Combining fuzzy and rough sets gives a key route in reasoning with uncertainty for real-valued data. Fuzzy rough sets encapsulate the related but distinct concepts of vagueness (for fuzzy sets) and indiscernibility (for rough sets), both of which occur as a result of uncertainty in knowledge. Vagueness arises due to a lack of sharp distinctions or boundaries in the data itself. This is typical of human communication and reasoning.

Fuzzy rough set concept has been implemented to surpass the deficiencies of the classical rough set approach in various aspects. The concept of a dependency function in a traditional rough set model into the fuzzy occurrence was proposed by Jensen and Shen [39; 40; 41; 42; 43] and introduced a feature selection algorithm using fuzzy rough set concept. This concept was extended by many researchers in their research articles. Another interesting mathematical framework to deal with imprecise and/or imperfect information is called the intuitionistic fuzzy (IF) set initiated by Atanassov [1; 2; 3]. An IF set is inherently considered as an extension of fuzzy set, which is determined by a pair of membership and non-membership functions. A fuzzy set provides a degree of belongingness of an element to a set while IF set gives a degree to which an element belongs to set as well as a degree to which an object does not belong to a set. The membership and non-membership values incite indeterminacy index, which produces the hesitancy in determining the degree to which an object meets a specific property. So, intuitionistic fuzzy set handles the uncertainty in a better way when compared to fuzzy set, as it handles latter uncertainty. In recent years, IF set theory has been effectively applied in the field of pattern recognition, decision analysis, medical image processing, etc. In spite of the fact that rough sets and IF sets both capture specific aspects of the same idea-imprecision, the combination of IF set theory and rough set theory are rarely discussed by the researchers.

In this thesis, we have proposed various intuitionistic fuzzy rough set models and presented different feature selection techniques using these models. All the proposed feature selection techniques are based on dependency function concept. This thesis consists of six chapters.

In the first chapter, we introduce the general concepts of feature selection and present the various techniques of feature selection. Moreover, preliminaries regarding feature selection has been given and work related to our thesis has been presented.

In the second chapter, we present a novel mechanism of attribute selection using tolerance-based intuitionistic fuzzy rough set theory. For this, we present tolerance-based intuitionistic fuzzy lower and upper approximations and formulate a degree of dependency of decision features over the set of conditional features. Moreover, the basic results on lower and upper approximations based on rough sets are extended for intuitionistic fuzzy rough sets and analogous results are established. In the end, the proposed algorithm is applied to an example data set and the comparison between tolerance-based fuzzy rough and intuitionistic fuzzy rough sets approaches for feature selection is presented. The proposed concept is found to be better performing in the form of selected attributes. The entire chapter in the form of a paper has been published in the journal of Expert System with Applications, 101, (2018) 205-212.

Third chapter defines the intuitionistic fuzzy decision of a sample and characterize its intuitionistic fuzzy information granule by introducing parameterized intuitionistic fuzzy relation. The proposed model is the generalization of fuzzy neighborhood rough set concept, which overcomes the drawback of fuzzy rough set approach by introducing a nearest neighbor of a sample to different intuitionistic fuzzy decision classes. We calculate the reduct of an information system by using degree of dependency approach. Furthermore, we present an attribute reduction algorithm for better understanding of our model. Finally, an illustrative example has been given to demonstrate our approach. The entire chapter in the form of a paper has been published in International Journal of Fuzzy System Applications, 7(2), (2018) 75-84.

In the fourth chapter, we have proposed a novel feature selection technique for partially labelled data set based on intuitionistic fuzzy rough set theory. Moreover, we have presented supporting theorems and proposed a novel algorithm to compute reduct based on our method. Finally, we have presented supremacy of our approach over fuzzy-rough technique by considering a partially labelled information system. The entire chapter in the form of a paper has been accepted in Advances in Intelligent Systems and Computing, Springer (2018).

Fifth chapter establishes a novel intuitionistic fuzzy rough set model by combining intuitionistic fuzzy and rough set with (α, β)-indiscernibility concept. Furthermore, we justify our proposed model by using supporting theorems. Moreover, a feature selection technique is proposed based on this model and a suitable algorithm is given to calculate the reduct for high dimensional datasets. Finally, an illustrative example is presented to demonstrate the effectiveness of our proposed concept.

Sixth chapter concludes the entire thesis and future directions has been presented.