8th ERES MEETING – ALICANTE, SPAIN - 24th - 26th JUNE 2001

"Within Real Estate Portfolio" Diversification through Rough Set Theory[1]

By
Maurizio d'Amato

Home Address: 272,Calefati St.- Bari - Italy
Home telephone: 39 80 5216009
e-mail : maurizio.damato@tiscalinet.it

Abstract

The work explores the opportunity to define a real property asset through rough set analysis proposed by Pawlak (Pawlak,1982), essentially in order to analyze information in uncertain context. The methodology was applied to property mass appraisal

(d'Amato, 2000). The procedure seems to be interesting in those markets where information is not easy to find because it let the "data speak on their own". The work is organized as follows, after a brief introduction, the first paragraph will offer a literature review concerning "within real estate" diversification techniques. In the second paragraph there will be a brief introduction to rough set analysis and in the third paragraph an application. Final remarks will conclude the work.

Introduction

The main goal of this work is trying to highlight the opportunity of applying Rough Set Theory in order to define groups of property assets for diversification strategies inside "within real estate" portfolios.

This theory based on two seminal works written by Pawlak[2] allows us to understand a problem through the observation of empirical data. In this work the methodology will be used for defining property assets without using statistical tools.

In a previous work this theory has been applied to mass appraisal[3]. The work is organized as follows. The next paragraph will offer a brief overview concerning "within real estate" portfolio literature. The following paragraph will offer a brief introduction to Pawlak's theory. The third paragraph will show a practical example. In the last paragraph will be given final remarks and future directions of research.

1. "Within Real Estate Portfolio" diversification: a brief literature review

The final goal of diversification is to reduce non systematic risk[4]. A "within real estate" portfolio is composed by real estate assets, whose nonsystematic risk is caused by several factors including lease terms, operating and financial leverage, tenant mix and location.

These factors are also linked to business cycles, secular trends, level of inflation and interest rates. "**Location"** and "**Property Types"** are the most important factors used as proxies to observe and measure risk factors. Several different contributions were given on the concept of location. One of the most important work using property type and geographic diversification was written by Miles and McCue[5] (1984). In this study the authors observed that diversification among real estate assets reduced unsystematic risk more than portfolio diversification did for common stock. Location has been considered either "homogeneous" or "functional". An example of "homogenous" region can be offered by NCREIF classification.

Economic diversification is the basis of economic interdependence. In order to reach this kind of diversification it is necessary to identify functional region. Since the 80's several works tried to define new methodologies in order to develop the concept of economic diversification. Hartzell, Hekman and Miles[6] (1986) suggested another approach to "within real estate" portfolio diversification taking into account MSA (metropolitan statistical area) growth, property type and lease maturity instead of the traditional property type and location. In the work they called for "more exacting categories" reliant on economic characteristics. The "Four Regions" model was improved by Hartzell Shulman and Wurtzbach[7] (1987) using eight economically homogenous regions. In a different way the concept behind this work was highlighted for the first time in his book by Joel Garreau[8]. This author divided in 9 fairly homogeneous zones basing the distinction on his experience as a newspaper reporter. Hartzell Shulman and Wurtzbach found that "…*regional diversification does matter for real estate portfolios, in the sense that the eight region categorization produces lower correlation coefficients than the traditional classification in four regions…*" and more extensively "*this study represents an attempt to move from mere geographical diversification to a more economic base – oriented concept*"[9]. Malizia and Sirmons[10] (1991) confirmed that economically based diversification strategies were superior because they were based on historical economic relationship rather than geopolitical boundaries. The analysis of specific economic drivers able to recognize metropolitan areas that allow an efficient diversification was developed by Mueller and Ziering[11] (1992). Hudson-Wilson[12] succeeded in defining these drivers using K-Means clustering algorithm on "derived market return". Risk and return were dependent on three variables: 1) Location 2) Property Types 3) Financial Structure. As a consequence several different property types were gathered together and the author says "…*This methodology hypothesizes that because two properties are located in different urban areas or are classed as different property types does not mean that each necessarily brings different risk and return characteristics to the portfolio*…". Goetzmann and Wachter (1995) showed a bootstrapping methodology for analyzing the exactness of clustering algorithm. Several works focused their attention on the concept of property type, Grissom Kuhle and Walther[13] (1987) used data from Houston and Austin for the years 1975 -1983 in order to observe the effect of diversification inside a "within real estate" portfolio. They concluded that "…*Effectively diversifying across either asset types or geographical locations leads to significantly lower amounts of unsystematic risk…*". Firstenberg , Ross and Zisler[14] (1988) confirmed the same conclusions observing that the "…*composition of a portfolio among geographic locations and property types can increase the investor's return for a given level of risk*…".

Mueller and Laposa[15] (1995) suggest to join property type diversification with another methodology.

A general overview concerning "within real estate" diversification strategies was offered in a recent work of Viezer[16](2000) in which thirteen different methodologies of diversification were compared. A real estate asset seems to be a complex mix among macroeconomic and microeconomic factors, property types and geographic features.

The main goal of this work is to offer a further possible methodology to define a real estate asset taking into account several diversification strategies in the same time.

2. A brief introduction to Rough Set Theory

In two famous works[17] Zdzislaw Pawlak introduced this theory in order to analyse information in uncertain context. In spite of common stock a property asset is not easy to define and the real estate market information is less available than stock market one. The complexity of the features of this kind of asset seems to suggest the use of a multicriteria methodology in uncertain context. Each element defined "object" inside a universe can be associated to information which relies on several attributes describing the element. In the case of a real estate asset it is possible to define as an object a property asset whose attributes could be property type, location etc. It is quite evident that real estate asset are more complex than stock. They involve a huge number of features. Property assets with similar attributes can be considered **indiscernible** at a certain information level.

Each property asset attribute can be "Certainly, Possibly and Certainly not". Furthermore, an indiscernible property asset is defined as an *elementary set.* A group of property assets as a subset of the universe of all properties can be defined as a union of two ordinary sets. The former is the so-called *"positive region*" or lower approximation, while the latter will be defined as a *"possible region*" or upward approximation. The rough set will be composed by these approximations[18]. The border between positive and possible regions is defined "*boundary region*" of "rough sets".

The positive of U set is composed by all the elements included in Y. On the other side the upward approximation is defined by the elementary sets with a non-empty intersection with U. The elements could belong to U or not. According to the information level it is not certain if some elements belong to U or not. An imprecise concept can be described with a couple of precise concepts: positive or possible regions.

The membership of a property asset to a cluster is described through lower or upward approximations.

In fact the property assets "…*belonging to the same category are not distinguishable…their membership status with respect to an arbitrary subset of the domain may not be clearly definable. This fact leads to the definition of a set in terms of lower and upper approximation…*"[19]

This process of knowledge allows us to discover causal relationships among the available data. In this methodology both qualitative and quantitative data can be used. There is not a preliminary analysis of the consistency of data because both bad and good data are useful and the “data speak for themselves”. The importance of the attributes is revealed by analysing the problem. The results will be defined through decision rules such as those based on “if…then” which will define the cluster and the kind of a particular real property asset.

The first step is the so-called “information table”. The row of this table will be filled with the elements. In this case the elements of the set are represented by the property assets to analyse. In the column the different attributes of the elements will be listed. Quantitative or qualitative evaluations of each attribute of the elements will be put inside the cells. The information table S is

$$\mathbf{S = < U, Q, V_q, f >} \quad \textit{(a)}$$

Where U is the universe or a finite set of elements (or real property assets), Q is a finite set of attributes or features (characteristics of real estate assets), V_q is the domain of attribute q and ***f*** is the so-called information function[20] that could be described as follows:

$$\mathbf{f : U \times Q \rightarrow V} \text{ and } f\ \mathbf{(x, q) \in V_q \quad \forall\ q \in Q} \text{ and } \mathbf{x \in U} \quad \textit{(b)}$$

Vectors will describe all the elements of U. This vector, also called description, will show the value that an attribute assumes for x inside Q set and it can be defined as Des_Q (x).

The object $x \in U$ will be described using a non-empty subset $P \subseteq Q$. For each subset of features P there is a indiscernible relation to U that could be indicated as Ip, where

$$\mathbf{Ip = \{(x, y) \in UxU: fq\ (x) = fq\ (y),\ q \in P\}}$$

This binary indiscernment[21] is an equivalent relationship. The couple (x, Ip) defines an *approximation space*. If $(x, y) \in Ip$, then x and y are P-indiscernible. Furthermore, if P= Q, the elementary Q sets are called atoms.

Another important concept is the union. It is possible to define *upward approximation* or $P_U X$ the subset of U composed by the elements belonging to P that have one element at least similar to X set. *Downward approximation* of X or $P_L X$ is the subset of U whose elements belong to P elementary set *included in X set,* and only to them.

The difference between these sets is a X boundary. It is defined as a BN_p (X) and it could be mathematically described as:

$$\mathbf{BN_p (X) = P_U X - P_L X}$$

If the frontier is empty, then X is the union of several ordinary sets defined through the union of several elementary P sets.

"*The lower approximation is a description of the domain objects which are known certainly as belonging to the interest subset, whereas the upper approximations"*

The methodology is suitable for a non-perfect information, a "granularity" of the information. Many dimensions influence the granularity of information such as: the quality of attributes, the number of attributes, and the domains of each attribute. The quality of results is strongly dependent on the information, on the capability to classify the information and on the level of confidence and knowledge of the problem.

A minimum subset of attribution (called "reduct") can be defined. This allows the analyst to have the same approximation of U of the complete set of features of P.

If $P \subseteq Q$ and $p \in P$ a feature is not important in P. P is defined orthogonal if all the attributes are important. P set is independent if all the attributes are important. The subset P' is a "reduct" of P if P' is independent and $I_p' = I_p$

In an information table there could be more than one "reduct" of P and the "*core*" of P is the set containing all the indispensable attributes of P. The core is inside each "reduct" of P which is considered as the most important subset of attributes of Q. No element of this sub set can be removed without diminishing the information quality.

An information table becomes a decision table if the attributes are divided in conditional attributes (C set) and decisional attributes (D set) showing the causal relationship between them. Decision rules are based on logic prepositions such as "if..then". The first part of the preposition is referred to one or more conditional attributes and the second part is represented by the decisional feature set.

There are two general types of decision rules. The former is the "exact decisional rule" or deterministic where the decisional set contains conditional attributes, the latter is the "approximated decision rule" in which some conditional attributes are included in the decisional set.

The granularity of the system could become higher when the information is based on few observations.

It is shown "…*By using this attribute it is possible to build a rule that classifies a given training set 100% correct; needless to say, the rule will not perform on an independent test set…*"[22]. Significant tests have been developed mainly based on randomization technique[23]. Furthermore a criterion for model selection based on minimum description length principal[24] defines the better selection of the model to explain the attribute.

The methodology introduced seems to be applied in order to identify a specific real property asset. As both qualitative and quantitative information is taken into account, several diversification methodologies can be used at the same time.

3. An Application

In order to show how this procedure works 10 real properties were taken into account[25]. A continuous scale was considered.[26]

The real estate assets were "approximated" through three attributes in order to classify the assets. Unfortunately this kind of study encountered great difficulties. The former is inside the market. In fact, Italian real estate market data are often unavailable. Few data were considered, but the main goal of this contribution is showing how the procedure works. In this way U set will be composed by ten elements, and the Q set is defined through the most important attributes chosen to classify the real property assets. Starting from the classification of data, decisional rules will be defined.

This method relies also on the principle of indifference[27] which states that in absence of further knowledge all basic events are "assumed to be equally likely"

Real Estate asset manager makes the delicate choice of features and their measures and if one or more important attributes are not considered, the results of the appraisal process will be not reliable.

A great help comes from several developed informatic tools that make the calculation and the definition of rules easy using a great amount of data[28].

Below , the so called "Information Table" is defined. Each real property is classified through four characteristics. The first is Geographic Location. The city of Bari was divided in “three concentric zones” Central, Semicentral, Peripheral. The second classification is based on a property type. Residential and Office property assets were essentially chosen. The REMO of Polytechnic of Bari collected these data. An economic base classification was impossible because of the lack of specific data. Return were calculated through the formula :

$$Ri_{(t)} = \frac{NOI + [\,Mvi(t\text{-}1) - Mvi(t)] + PS - CI}{Mvi(t) + \frac{1}{2}\,CI - \frac{1}{2}\,PS - \frac{1}{3}\,NOI}$$

Where **Ri(t)** il the holding period return, **Mvi(t)** the beginning of a period market value and the **Mvi(t-1)** the end of period market value, **CI** capital improvements , **NOI** net operating income and **PS** partial sales. Both the initial market value and the final one were estimated through a real estate exchange (Borse Immobiliari). At the present in Italy there are not specific analysis and the returns were estimated on an annual basis for a five-year holding period. Because of the lack of data estimating both return and their variability (as a risk measure) on more precise data is impossible. For this reason the main goal of this work is to highlight the contribution of Rough Set methodology to group real property asset inside a “within real estate” portfolio. A further analysis on more reliable data could test the reliability of the method. Table 1 shows the classification of the ten data “approximated” through three features:

Real Estate Asset	Geographic Location	Property Type	Return Profile
1	Central	Residential	0.041
2	Peripheral	Office	0.043
3	Central	Residential	0.044
4	Semi-central	Residential	0.039
5	Semi-central	Office	0.041
6	Semi-central	Residential	0.035
7	Central	Office	0.051
8	Peripheral	Office	0.041

9	Semi-central	Office	0.043
10	Central	Residential	0.045

Table 1- Information Table Referred to 10 real property

In this information table there are the following sets: U ={1,2,3,4,5,6,7,8,9,10 } Q = { GEOGRAPHIC LOCATION, PROPERTY TYPE, RETURN RISK RATIOS}. The table represents the information function f (x, q) which can be **exemplified** as f (5,PROPERTY TYPE) = OFFICE. The value of the attributes V_q will vary depending on several different scales. It is obvious that the third attribute could be also the ratio between risk and return.

The returns described in the table 1 can be grouped through Rough Set Theory Rules. It must be highlighted that this work does not provide a practical application of the method because of the lack of market data. The work offers a simple overview towards an application of this theory to " within real estate" market. From the information table a decision table can be shown:

Objects (U – SET)	**Conditional Attributes** (Q SET)		**Decisional Attributes** (d SET)
Real Estate Asset	**GEOGRAPHIC LOCATION**	**PROPERTY TYPE**	**RETURN**
1	Central	Residential	0.041
2	Peripheral	Office	0.043
3	Central	Residential	0.044
4	Semi-central	Residential	0.039
5	Semi-central	Office	0.041
6	Semi-central	Residential	0.035
7	Central	Office	0.051
8	Peripheral	Office	0.041
9	Semicentral	Office	0.043
10	Central	Residential	0.045

Table 3- Decision Table

Observing the data no specific rules is developed and the "data speak for themselves" Several equivalent classes will be developed using indiscernibility relationship. In order to analyse the data through the decisional attribute "Return" several classes were built:

Group	Central Residential, Peripheral Office,Central Residential, Semicentral Office, Peripheral Office, Semicentral Office, Central Residential (asset A)	Semicentral Residential **(asset B)**	Central Office **(asset C)**
Return Profile	**0,041 – 0,045**	**0,035- 0,039**	**0,051**

Table 4 – Assets classes

The attribute behaviour is always supposed to be the same inside the set. Table n. 5, indicated below, shows several groups separating the conditional attributes from the decisional one (return profile). In these classes of equivalence there are identical measures for the same attribute.

Conditional Features – Q	**Classes of Equivalence RelationshipIp**
í Geographic Locationý **íProperty Typeý** **íGeographic Location, Property Typeý**	{1,3,7,10} {2,8} {4,5,6,9} {1,3,4,6,10,} {2,5,7,8,9} {1,10,3} {2,8} {4,6}{5,9}
Decisional Feature – Return Profile – d	**Classes of Equivalence RelationshipId**
íReturn Profileý	{1,2,3,5,8,9,10}{6,4}{7}

Table 5- Defining classes of equivalence

These classes allow the asset manager to develop several “if…then” rules whose final issue is to classify both the property assets taken into account and to develop a specific framework to analyse other assets. It must be highlighted that this methodology can be applied to the results of several diversification methodologies. In the same time, developing “if...then” rules based on geographic, economic and property type methodologies is possible in order to define the return profile of a property asset. According to the final issue of the work the rule should be based on “deterministic” rules instead of “non-deterministic rules”. The former rule defines a stronger causal relationship between attributes and return profile. This specific rule is based on downward approximation and it will be defined through the following relationship

(X, d) Î Q ->d ⇔ x Ç d ¹ 0

Q -> d

Q => d if and only if I_q Í I_d

As a consequence the relationships **Q -> d** will be

íGeographic Locationý → Return

{1,3,7,10}{1,2,3,5,8,9,10}

{1,3,7,10}{6,4}

{1,3,7,10} {7}

í2,8ý í1,2,3,5,8,9,10ý

{4,5,6,9}{1,2,3,5,8,9,10}

{4,5,6,9}{6,4}

íProperty Typeý → Return

{1,3,4,6,10,}{1,2,3,5,8,9,10}

{1,3,4,6,10,}{6,4}

{2,5,7,8,9}{1,2,3,5,8,9,10}

{2,5,7,8,9}{7}

íGeographic Location, Property Typeý → Price

í1,10,3ý í1,2,3,5,8,9,10ý

í2,8ý í1,2,3,5,8,9,10ý

í4,6ý í6,4ý

í5,9ý í1,2,3,5,8,9,10ý

This kind of relationship (downward approximation) is verified in the following cases:

Q => d	**Deterministic Rules**
í GEOGRAPHIC LOCATIONý → íRETURNý	**í2,8ý í1,2,3,5,8,9,10ý**
íPROPERTY TYPEý → íRETURNý	**0**
{GEOGRAPHIC LOCATION,PROPERTY TYPE} →{RETURN}	**í1,10,3ý í1,2,3,5,8,9,10ý**
	í2,8ý í1,2,3,5,8,9,10ý
	í4,6ý í6,4ý
	í5,9ý í1,2,3,5,8,9,10ý

Therefore the following rule based on the observation of data will be created:

Geographic Location → Return

IF *Peripheral Office* THEN *asset return class A*

Property Type → Return

NON-DETERMINISTIC RULES

Geographic Location and Property Type → Return

IF *Residential* AND *Central* THEN *Return Asset Class* will be A

IF *Peripheral* AND *Office* THEN *Return Asset Class* will be A

IF *Semicentral* AND *Residential* THEN *Return Asset Class* will be B

IF *Semicentral* AND *Office* THEN *Return Asset Class* will be A

As it is possible to observe for the asset class return C there is not specific rule. In fact a rule is based on empirical observations that are available only for asset class return B and A. Another important remark is that the asset class can be made in several ways including risk return ratios. An increasing number of information will allow the analyst to define better the property asset characteristics.

The methodology puts several different diversification strategies working together and there are not specific assumptions as in regression analysis. In this method property asset features are analysed relying only on the principle of indifference. There is a bottom-up process starting from the full attribute set trying to reduce it in a few deterministic rules. While in the statistical model there are few variables for many requested observations, in Rough set theory it is possible to consider a huge number of features with few data[29].

The link between property features and its return can be pointed out without defining a specific model using both qualitative and quantitative variables. If a huge amount of data and features will be analysed then rules will be more and more sophisticated. An important element to observe is the "strength of a rule"[30] defined by counting the number of objects which a rule refers to.

Final Remarks and Future Directions of Research

At the end of this work it is possible to highlight some results.

- The work is not focused on testing the methodology but the final issue is showing how it can be applied to "within real" estate portfolio.
- This procedure allows more than one diversification methodologies to work together
- The results will be used either to classify a property asset or to foresee the future return profile of assets.
- The work can be improved defining return and risk ratio profile. In Italy a strictly application of the methodology seems to be impossible because of the lack of official data and statistics.

A future direction of research can be a comparison between this methodology with another one. A comparison between statistical tools and Rough Set Theory could be suggested.

References

[1] The work was developed in the Faculty of Economics in Alicante during a research program financed by CRi.SMI and the Italian National Council of Research in September 2000. I would like to thank you Paloma Taltavul de la Paz for suggestions and help

[2] See Pawlak Z. (1982), *Rough Sets*, International Journal of Information and Computer Sciences, 11,341-356 and Pawlak Z. (1991), *Rough Sets. Theoretical Aspects of Reasoning about Data*, Kluwer Academic Publisher, Dordecht

[3] M. d'Amato (2000), *Appraising Properties with Rough Set Theory*, Refereed Paper Accepted by the 7th Pacific RIM Congress Adelaide 21-24 January 2001

[4] Harry Markowitz (1959), *Portfolio Selection*, New Haven, Conneticut, Yale University Press.

[5] M.E. Miles and T.E.McCue (1984) , *Commercial Real Estate Returns* , AREUEA Journal, Fall

[6] Hartzell D.J., Hekman J.S. and Miles E.M. (1986), *Diversification Categories in Investment Real Estate* , Journal of the American Real Estate and Urban Economics Association, 14:2, 230-254

[7] Hartzell D.J., Shulman D. and Wurtzebach C.H. (1987), *Refining the Analysis of Regional Diversification for Income Producing Real Estate*, Journal of Real Estate Research, 2:2, 85-95

[8] J.Garreau (1981), *The Nine Nations of North America*, New York: Avon Books

[9] Hartzell D.J., Shulman D. and Wurtzebach C.H. (1987), Ibid., p.94

[10] Malizia E.E. and R.A.Simons, *Comparing Regional Classifications for Real Estate Portfolio Diversification*, Journal of Real Estate Research,1991, 6:1,53-77

[11] G.R.Mueller and B.A.Ziering (1992), *Real Estate Portofolio Using Economic Diversification*, Journal of Real Estate Research, 7:4, 375-386

[12] Hudson-Wilson S.(1990), *New Trends in Portfolio Theory*, Journal of Property Management, 55, 57-8 and Hudson-Wilson S., *Sampco: A Hypothetical Portfolio Analysis*, in *Managing Real Estates Portfolio*, S.Hudson Wilson and C.H.Wurtzebach (Eds.) Burr Ridge: Irwin Professional Publishing

[13] T.V. Grissom, J.L.Kuhle, and C.H.Walther (1987), *Diversification Works in Real Estate,Too*, Journal of Portfolio Management, 13:2 , 66-71

[14] P.M. Firstenberg, S.A. Ross, and R.C. Zisler (1988), *Real Estate: The Whole Story*, The Journal of Portfolio Management, 14:3, 22-34

[15] Mueller G.R. and S. P. Laposa (1995), *Property Type Diversification in Real Estate Portfolios: A Size and Return Perspective*, Journal of Real Estate Portfolio Management, 1:1, 39-50

[16] T.W.Viezer (2000), *Evaluating "Within Real Estate" Diversification Strategies*, Journal of Real Estate Portfolio Management, 6:1 , 75-95

[17] See Pawlak Z. (1982), *Rough Sets*, International Journal of Information and Computer Sciences, 11,341-356 and Pawlak Z. (1991), *Rough Sets. Theoretical Aspects of Reasoning about Data*, Kluwer Academic Publisher, Dordecht

[18] Note that these approximation sets will be one in case of ordinary sets

[19] W.Ziarko (1993), *A Brief Introduction to Rough Sets, The First International Workshopon Rough Sets: States of Art and Perspectives*, University of ReginaSaskatchewan

[20] See Pawlak, Z. (1991) *Rough Sets. Theoretical Aspects of Reasoning About Data*. Dordrecht: Kluwer Academic Publishers

[21] In some later versions of this methodology a dominant relationship was used. It allows to use some attributes with preference-ordered scales or criteria. See Greco S., Matarazzo B. and Slowinsky R. (1998) " *A new rough set approach to evaluation of bankruptcy risk*" in Operational Tools in Management of Financial Risk, edited by C. Zopounidis, pp.121-136, Dordrecht, Boston, Kluwer Academic Publishers

[22] Holte R.C. (1993) *Very Simple Classification Rules Perform Well On Most Commonly Used Datasets,* Machine Learning, 11:63-91

[23] Duntsch I and Gediga G. (1997) *Statistical Evaluation of Rough Set Dependency Analysis*, International Journal of Human-Computer Studies, 46 589-604

[24] For everything see the two works by Rissanen J. (1985), *Minimum Description Length Principle*, in S. Koltzand N.L. Johnson editors Encyclopedia of Statistical Sciences,523-527 N.Y. Wiley; Rissanen J. (1978) *Modeling by the Shortest Data Description*, Automatica, 14, 465-471

[25] These properties were given by the Real Estate Market Observatory of 1st School of Engineering of Polytechnic of Bari inside the Architecture and Planning Dept.

[26] Rough Set works with a continuous scale, too. See the works by Browne C. Duntsch I and Gediga G. (1998). *IRIS revisited: A comparison of Discriminant and Enhanced Rough Set Data Analysis* in Slowinski R. editor (1992) *Intelligent Decision Support: handbook of Applications and Advances of Rough Set Theory* Volume 11 System Theory, Knowledge Engineering and Problem Solving, Kluwer Dordrecht pages 345-368

[27] See Bacchus F, Grove A. J. Halpern J.Y., Koller D. (1994) *From Statistical Knowledge Bases to Degrees of Belief.* Technical Report 9855, IBM

[28] See for example ROSETTA software inside the following web-sites: **www.idt.unit.no/~aleks/rosetta/rosetta.html**

[29] For an analysis of the differences between statistical model and Rough Set see Duntsch I. & Gediga G. (1997), *Roughian – Rough Set Information Analysis*, In Sydow A. Editor (1997), Proc.15th IMACS World Congress Vol.4, Berlin, Wissenschaft und Technik Verlag

[30] Krusinka E., Babic A., Slowinski R & Stefanowski J.(1992), *Comparison of the Rough Set Approach and Probabilistic Data Analysis Techniques on a common set of medical data*, In Slowinski R., (ed)(1992), *Intelligent Decision Support.Handbook of Applications and Advances of the Rough Set Theory*,Dordrecht Boston Kluwer Academic Publishers