

IMPLEMENTASI *RANKER PRINCIPAL COMPONENT* DAN *METODE ASSOCIATION RULE* DALAM MENENTUKAN FAKTOR EKONOMI DI ASEAN

Fajri Rakhmat Umbara, Ridwan Ilyas

Jurusan Informatika

Universitas Jenderal Achmad Yani

Cimahi

fajri.rakhmat@lecture.njani.ac.id, ilyas@lecture.unjani.ac.id

Abstrak

Kondisi perekonomian suatu negara sangat berpengaruh bagi kelangsungan hidup Negara tersebut. Seperti contoh yang terjadi pada Negara Yunani beberapa tahun yang lalu dimana mengalami krisis ekonomi yang tidak terelakkan. Kondisi perekonomian tersebut tidak terlepas dari faktor – faktor ekonomi yang saling terkait didalamnya dan yang memang mempengaruhi Negara yang berkaitan, misalnya Produk Domestik Bruto (PDB), Nilai Tukar Mata Uang, Penghasilan Energi Baru Terbarukan dan lain sebagainya. Faktor – faktor itu yang menjadi tolak ukur apakah perekonomian suatu Negara dapat dikategorikan berhasil atau malah dapat menyebabkan krisis ekonomi. Fokus penelitian ini adalah untuk melihat dan mengevaluasi faktor ekonomi apa sajakah yang dapat mempengaruhi kondisi suatu bangsa, khususnya untuk Negara – Negara di ASEAN, termasuk diantaranya adalah Indonesia menggunakan seleksi atribut *Ranker Principal Component* dan metode *Association Rule*. Data ekonomi didapat dari data yang disediakan oleh World Bank selama 50 tahun lamanya, atau antara tahun 1967 hingga tahun 2016 yang terdiri dari sekitar 1400 atribut / faktor ekonomi untuk 10 negara ASEAN yang terdiri dari Indonesia, Thailand, Malaysia, Singapura, Filipina, Brunei Darussalam, Vietnam, Laos, Myanmar, dan Kamboja.

Kata kunci :

Faktor ekonomi asean, ranker principle component, association rule

Abstract

Economic condition may cause a balance for every country in the world. For example, Greece has been collapse after economic crisis in 2008. This phenomenon caused by economic factors that

influence and depends on each other such as GDP, exchange rate, energy production, and so on. These factors become a benchmark for countrys that is can cause crisis or not. This research try to searching and evaluating these factors which influence head to the crisis in ASEAN, include Indonesia using atribut selection methods : ranker principle component and association rules methods. The economic data are downloaded from World Bank site and has been collected for 50 years between 1967 and 2016 which has 1400 attributes or economic factors for 10 ASEAN countrys such as Indonesia, Thailand, Malaysia, Singapore, Brunei Darussalam, Vietnam, Laos, Myanmar and Cambodia.

Keywords :

ASEAN economic factors, ranker principle component, association rule

I. PENDAHULUAN

I.1 Latar Belakang

Eksistensi suatu Negara, salah satunya dapat dilihat dari kondisi perekonomian Negara tersebut. Beberapa tahun yang lalu, tepatnya pada tahun 2008, Negara – Negara di Eropa mengalami krisis ekonomi yang mana kondisi tersebut mempengaruhi banyak Negara yang posisinya bahkan diluar Eropa itu sendiri. Negara Yunani adalah contoh bagaimana dampak krisis ekonomi di Eropa, dimana Negara ini sulit bertahan karena hanya mengandalkan mata pencaharian di sektor pariwisata sebagai usaha untuk mempertahankan kondisi perekonomiannya. Sektor pariwisata ini hanya mempengaruhi nilai Produk Domestik Bruto dari Negara Yunani ini, padahal masih banyak faktor ekonomi lain yang bisa

ditingkatkan nilainya agar Negara tersebut mata pencahariannya tidak bergantung hanya pada satu sektor saja.

Indonesia, sebagai salah satu Negara anggota ASEAN juga pernah mengalami krisis ekonomi pada tahun 1997 – 1998. Faktor ekonomi yang sangat mempengaruhi pada saat itu adalah Produk Domestik Bruto dan Hutang Luar Negeri, dimana nilai Hutang Luar Negeri lebih tinggi daripada Produk Domestik Bruto.

Untuk prediksi krisis ekonomi di Indonesia sendiri pernah dilakukan pada penelitian terdahulu, dengan memperhitungkan faktor ekonomi Produk Domestik Bruto dan Hutang luar negeri, maka Indonesia akan mengalami krisis kembali pada tahun 2030 apabila kondisi dari kedua faktor tersebut tidak diperbaiki (F. R. Umbara 2016).

Penelitian lain yang menunjukkan kaitan antara faktor ekonomi dengan krisis juga pernah dilakukan (G. Kaminsky, S. Lizondo dan C. Reinhart. 2008), sebagian besar negara dunia dipengaruhi oleh salah satu faktor yaitu *Gross Domestic Product* (GDP).

Untuk negara D8 (S. Razzaqi et al. 2011) masalah terbesar dari penyebab ketidak seimbangan ekonomi adalah masalah *Energy*.

I.2 Rumusan Masalah

Berdasarkan latar belakang diatas maka dirumuskan permasalahan pokok yaitu menemukan dan menganalisis faktor ekonomi di asean dengan menggunakan seleksi atribut *Ranker Principal Component* dan metode *Association Rule*.

I.3 Batasan Masalah

Adapun batasan masalah dalam penelitian ini antara lain :

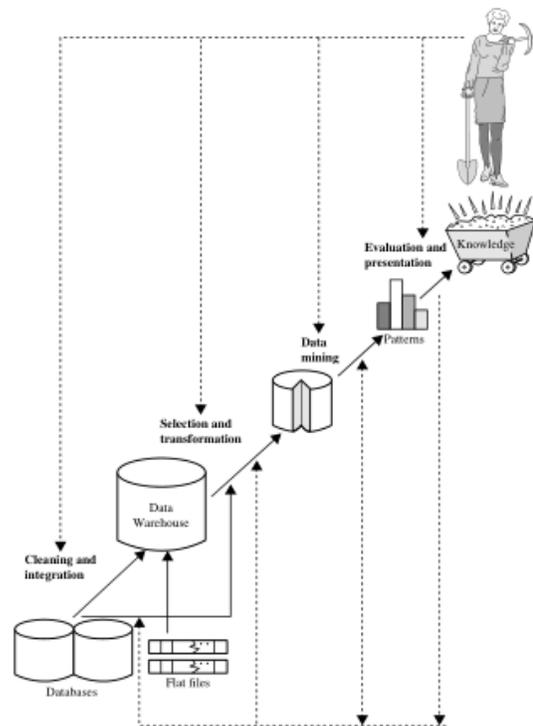
- Faktor - faktor ekonomi yang di evaluasi hanya pada 10 negara ASEAN saja
- Tools yang digunakan dalam penelitian ini adalah WEKA 3.8 dengan nilai masukan dari tools adalah *default value*
- Pengisian *missing value* hanya menggunakan nilai *mean* atau *modes* saja

II. KAJIAN LITERATUR

II.1 Data Mining

Data Mining merupakan ilmu yang berkembang untuk melakukan penggalian terhadap data, dimana sesungguhnya dalam kumpulan data terdapat suatu informasi yang bisa digali lebih dalam dengan hasil yang lebih baik daripada mencari informasi hanya dengan menggunakan suatu *Query* saja (J. Han, Micheline Kamber, & Jian Pei. 2012).

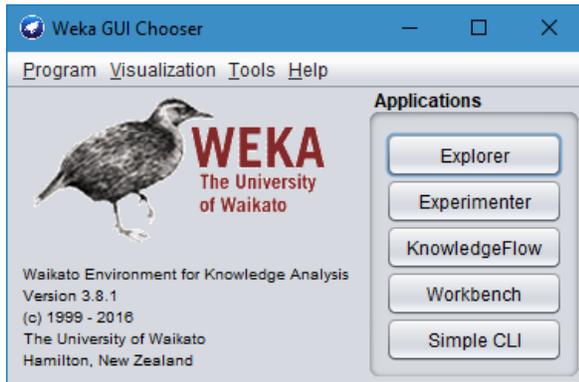
Data mining juga dikenal dengan istilah KDD (*Knowledge Discovery and Data*), dimana untuk menghasilkan suatu pengetahuan yang berguna, maka tahapan yang harus dilakukan adalah seperti gambar dibawah ini :



Gambar 1. Data Mining dalam Penggalian data KDD (Sumber gambar : J. Han, Micheline Kamber, & Jian Pei. 2012)

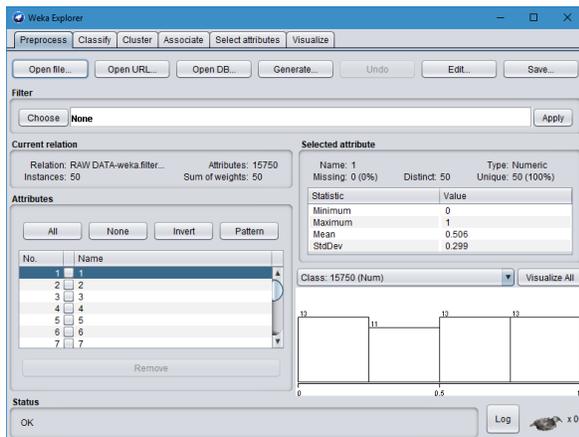
II.2 WEKA

WEKA (Waikato environment for Knowledge Analysis) adalah mesin untuk mengolah data yang dibangun menggunakan Bahasa java dan bersifat *open source*. WEKA digunakan untuk analisis data dan sebagai salah satu *tools machine learning* yang cukup tangguh (S. Singhal & Monika Jena. 2013).



Gambar 2. WEKA Tools

WEKA juga memberikan penyajian GUI yang mudah digunakan dan cepat, serta hasil yang langsung dapat digunakan dan dianalisa. Dengan menggunakan format file data berupa ARFF, maka untuk tiap atribut dapat menyimpan nilai dan jenis daripada atribut itu sendiri, misalnya nominal, biner, ordinal ataupun numerik (I. H. Witten, Eibe Frank, & Mark A Hall. 2011).



Gambar 3. Tampilan GUI WEKA

II.3 Missing Value

Missing value adalah suatu kondisi dimana suatu atribut tidak memiliki nilai atau nilai yang diinginkan tidak sesuai, sehingga atribut tersebut harus ditambahkan dan diubah sedemikian rupa agar dapat digunakan, namun pengisian *missing value* tersebut tidak mengubah makna daripada data yang disajikan untuk dianalisa (M. L. Brown & John F Kross. 2003).

Di beberapa kondisi, *missing value* dapat menjadi *noise* apabila jumlahnya dalam suatu kumpulan data sangat banyak. *missing value* bisa dikonstruksi dari hal yang sederhana, misalnya mengisi nilai dengan nilai *mean*, *median*, atau *modes* saja (J. Han, Micheline Kamber, & Jian Pei. 2012). Selain itu, ada beberapa cara lain yang bisa dilakukan dalam menanggulangi *missing value*, diantaranya adalah :

1. *Tuple* tidak digunakan dalam *dataset*
2. Dimasukkan secara manual
3. Gunakan nilai global untuk mengisi *missing value*
4. Gunakan nilai sampling dari kelas yang sama
5. Gunakan nilai yang mungkin.

II.4 Normalisasi

Normalisasi digunakan untuk membuat semua nilai menjadi satu nilai yang homogen. Misalnya dalam suatu kondisi, data memiliki banyak atribut dengan banyak pula nilai yang berbeda, mulai dari type, jenis, ataupun satuan yang digunakan. Ketidakteragaman nilai dapat menyebabkan metode yang digunakan untuk melakukan analisis data atau data mining akan mengeluarkan hasil yang tidak sesuai, sehingga kesimpulan yang dihasilkan juga akan memiliki tingkat kepercayaan dan tingkat kebenaran yang rendah.

Normalisasi adalah tahapan dari Data Mining yaitu Transformasi Data (J. Han, Micheline Kamber, & Jian Pei. 2012) dimana terjadi perubahan dari bentuk data namun tidak merubah nilainya.

Dalam penelitian ini, normalisasi yang dilakukan adalah memiliki rentang nilai antara 0 dan 1.

II.5 Ranker Principle Component

Ranker Principal Component merupakan *unsupervised attribute evaluator*, dimana menggunakan *Principal Component Analysis* (PCA) dan *Ranker Searching* untuk perankingan atribut,

mulai dari nilai yang lebih besar dianggap sebagai atribut paling penting hingga nilai yang terkecil yang dianggap sebagai atribut sangat tidak penting. PCA sendiri merupakan metode yang banyak digunakan untuk menganalisa pola konsumen pada penelitian yang lain (F. Westad, M. Hersleth, P. Lea, H. Martens. 2003).

Langkah – langkah perhitungan *Ranker Principal Component* adalah sebagai berikut :

1. Hitung rata – rata matriks dari data yang ada
2. Hitung ko-varian matriks
3. Hitung nilai Eigen dan Vektor Eigen
4. Urutkan Vektor Eigen dari nilai tertinggi sampai terkecil
5. Pilih Vektor Eigen terbaik, kemudian disebut *Eigen Face*
6. Gunakan nilai *Eigen Face* ke algoritma *Ranker Searching*, dimana nilai atribut akan dievaluasi dan diurutkan berdasarkan nilai tertinggi sampai terkecil.

II.6 Association Rule

Association rule, dalam penggunaannya di *data mining*, digunakan untuk melihat keterkaitan suatu atribut dengan atribut lain dalam bentuk *rule* atau aturan. Saat menggunakan metode ini terlebih dahulu mencari nilai *support* dan *confidence* dengan persamaan dibawah ini (dengan metode Apriori):

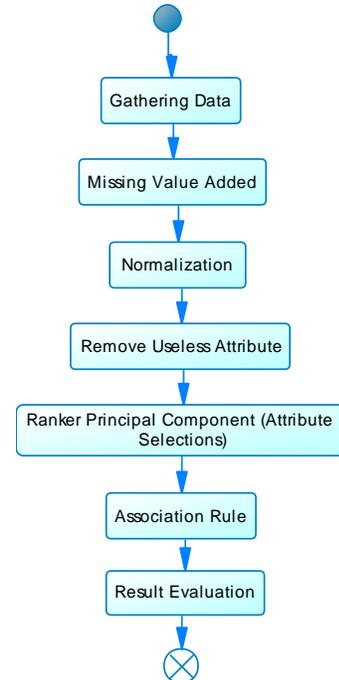
- a. $Support (A) = \frac{\text{Jumlah transaksi yang mengandung } A}{\text{Total transaksi}}$
- b. $Support (A \cap B) = \frac{\text{Jumlah transaksi yang mengandung } A \text{ dan } B}{\text{Total transaksi}}$
- c. $Confidence = P (A|B) = \frac{\text{Jumlah transaksi yang mengandung } A \text{ dan } B}{\text{Jumlah transaksi yang mengandung } A}$

Association rule dapat menemukan hubungan yang tepat antara satu atribut dengan semua atribut lain (D. Prayitno, 2015) dimana keterkaitan antara atribut ini menunjukkan mekanisme ketergantungan, mungkin dapat pula menunjukkan ketergantungan sebab dan akibat.

Rule yang terbaik adalah rule yang memiliki nilai *confidence* tinggi.

II.7 Metode Penelitian

Tahapan penelitian yang telah dilaksanakan dalam penelitian ini adalah sebagai berikut :



Gambar 4. Tahapan Penelitian

Langkah-langkah penelitian seperti yang digambarkan pada gambar diatas adalah langkah yang dilaksanakan untuk keseluruhan penelitian, sebagai berikut:

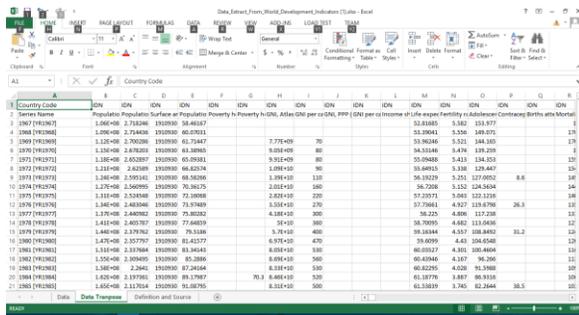
1. *Gathering Data*; melakukan pengambilan data faktor – faktor ekonomi di 10 negara ASEAN selama 50 tahun dari tahun 1967 hingga 2016 melalui situs World Bank Repository.
2. *Missing Value Added*; banyak terdapat missing value pada data yang diambil, untuk itu data dilengkapi dengan mengisi data yang kosong menggunakan nilai mean dari keseluruhan data.
3. *Normalization*; karena nilai atribut yang didapat bervariasi, maka dilakukan normalisasi nilai untuk keseluruhan atribut.
4. *Remove Useless Attribute*; banyak juga atribut yang tidak terpakai karena data yang sangat tidak lengkap bahkan kosong, untuk itu atribut dengan kondisi seperti ini dihilangkan.
5. *Ranker Principal Component Attribute Selection; unsupervised attribute evaluation*, akan

mengurutkan atribut dengan nilai kepentingan tertinggi akan diurutkan paling depan dan yang kepentingannya paling rendah akan diurutkan di posisi akhir.

6. *Association Rule*; untuk membuat model berupa rule hubungan antar faktor – faktor ekonomi di ASEAN.
7. *Result Evaluation*; menganalisis dan membuat kesimpulan terhadap hasil mode dari *association rule*.

III. ANALISIS DAN PERANCANGAN

Dataset menggunakan dataset yang diambil dari World Bank, pada situsnya : databank.worldbank.org, menggunakan dataset ekonomi ASEAN mulai dari tahun 1967 sampai dengan 2016 dengan format CSV seperti gambar dibawah ini :



Gambar 5. Dataset dengan vortmat CSV

Dalam penggunaan dataset tersebut, ada beberapa negara yang nilai atributnya kosong, untuk mengisinya digunakan nilai *mean* dari seluruh nilai atribut itu sendiri, sebahai contoh untuk nilai dari atribut *Urban Population*, atribut yang berisi jumlah penduduk yang tinggal di daerah urban, ada beberapa negara yang datanya tidak ada, mka diisi dengan nilai mean. Namun ada juga yang semua negara terdapat atribut yang nilainya *null*, maka atribut tersebut tidak digunakan.

Setelah dataset bersih, format CSV diubah menjadi format WEKA, yaitu ARFF agar data untuk tiap tiap atribut memiliki bentuk dan format yang diinginkan. Data dengan Format ARFF dilakukan normalisasi, lalu hasilnya kemudian dilakukan proses pengurutan atribut menggunakan *Ranker Principal Component*.

Hasil dari *Ranker Principal Component*, diperoleh table yang berisi seluruh faktor ekonomi yang memiliki nilai rangking namun sudah diurutkan dari paling tinggi hingga paling rendah. Dalam percobaan kali ini, diambil hanya 10 urutan terdepan dari hasil yang telah ada untuk melihat relasi antara kesepuluh atribut ini menggunakan metode *Association Rule*.

Adapun hasil pengurutan atribut yang sudah diubah bentuknya sedemikian rupa kedalam bentuk tabel agar dapat dibuat rulenya dengan menggunakan metode *Association Rule* yang ada pada WEKA seperti dibawah ini :

Tabel 1. Data Masukan *Association Rule*

Country	Population total	Population growth (annual %)	Surface area (sq. km)	Population density (people per sq. km of land area)	Poverty headcount ratio at national poverty lines (% of population)	Poverty headcount ratio at \$1.90 a day (2011 PPP) (% of population)	GNI, Atlas method (current US\$)	GNI per capita, Atlas method (current US\$)	GNI, PPP (current international \$)	GNI per capita, PPP (current international \$)
BRN	yes	yes	yes	no	no	no	yes	yes	yes	yes
IDN	yes	yes	no	no	no	yes	no	yes	no	yes
KHM	yes	yes	yes	no	yes	no	no	yes	no	no
LAO	yes	yes	yes	yes	yes	yes	no	yes	yes	yes
MMR	yes	yes	yes	yes	yes	no	no	yes	no	yes
MYS	yes	yes	yes	yes	yes	yes	no	yes	yes	yes
PHL	yes	yes	yes	yes	no	no	no	yes	yes	no
SGP	yes	yes	no	no	yes	yes	no	yes	no	yes
THA	yes	yes	yes	no	yes	yes	yes	yes	yes	yes
VNM	yes	yes	yes	no	no	yes	yes	no	no	yes

Di bagian baris menunjukkan 10 Negara ASEAN, dan kolom menunjukkan faktor ekonomi apa saja yang diambil. Nilai yang diisi untuk tiap tiap baris dan kolom adalah yes dan no, dimana ini menunjukkan apakah dengan urutan atribut tersebut, negara yang berkaitan memiliki keterurutan yang sama atau tidak.

Data pada Tabel 1 diolah menggunakan metode *Association Rule* untuk melihat keterkaitan antara atribut dari faktor ekonomi yang sudah terurutkan tadi, karena untuk masing – masing negara memiliki urutan faktor ekonomi yang berbeda – beda. Hasil olah data dengan metode *association rule* menghasilkan 10 buah *rule* terbaik yang memiliki nilai *minimum support* = 0.9 seperti dibawah ini :

1. *Population Growth* → *Populations* (*confidence* = 1)
2. *Population* → *Population* (*confidence* = 1)
3. *GNI per Capita* → *Population* (*confidence* = 1)
4. *GNI per Capita* → *Population Growth* (*confidence* = 1)

5. *Population Growth, GNI per Capita* → *Population* (confidence = 1)
6. *Population, GNI per Capita* → *Population Growth* (confidence = 1)
7. *GNI per Capita* → *Population, Population Growth* (confidence = 1)
8. *Population* → *GNI per Capita* (confidence = 0.9)
9. *Population Growth* → *GNI per Capita* (confidence = 0.9)
10. *Population, Population Growth* → *GNI per Capita* (confidence = 0.9)

Dari 10 rules yang dibentuk, terdapat 7 rule dengan nilai confidence = 1 dan 3 rules yang memiliki nilai confidence = 0.9. Ini berarti rule yang akan digunakan untuk kesimpulan sudah memiliki tingkat hubungan yang cukup tinggi.

Dari urutan atribut yang dihasilkan dari metode Ranker Principal Component, sudah terlihat atribut mana saja yang mempengaruhi dan yang memiliki urutan paling penting, namun setiap negara memiliki urutan yang berbeda – beda. Association rule digunakan untuk membuat kesimpulan dan melihat keterkaitan antar atribut yang berpengaruh untuk semua negara ASEAN. Dari Rules yang terbentuk, terlihat bahwa *Population* (atribut yang berisi nilai jumlah populasi tiap negara per tahun), *Population Growth* (atribut yang berisi nilai presentase pertumbuhan penduduk tiap tahun), dan *GNI per Capita* (atribut yang berisi nilai pendapatan nasional bruto per capita tiap negara per tahun) menjadi 3 atribut utama yang mempengaruhi kondisi ekonomi di negara – negara ASEAN.

IV. KESIMPULAN DAN SARAN

Dari 10 atribut yang dievaluasi dan dibentuk rule atau model menggunakan Association Rule, dapat dilihat bahwa jumlah populasi (*Populations, Total*) dan pertumbuhan populasi (*Populations Growth*) di negara – negara asean menjadi faktor ekonomi yang sangat berpengaruh. Kedua faktor tersebut saling mempengaruhi dan memiliki nilai Confidence yang tinggi pula (confidence = 1). Untuk itu, jumlah populasi dan pertumbuhan populasi dapat dijadikan suatu leading indicator di negara – negara ASEAN sebagai antisipasi datangnya krisis ekonomi.

Kajian diatas tentunya dikaji dengan menggunakan metode sains dan hasilnya bisa diperkuat dengan menggunakan analisis lain di bidang ekonomi sehingga kesimpulan terhadap faktor ekonomi apa saja yang mempengaruhi kondisi negara – negara di ASEAN akan lebih kuat.

REFERENSI

- D. Prayitno. 2015. Penerapan Association Rule Mining Pada data Bursa Efek Indonesia Dengan Algoritma Apriori Untuk Memilih saham Unggulan. *Jurnal Teknologi*, vol. 8, no. 2, pp. 96 - 103.
- F. R. Umbara. 2016. Prediksi Jangka Panjang Krisis Ekonomi Di Indonesia Menggunakan Model Sistem Dinamik Dioptimalisasi Menggunakan Algoritma Genetika Secara Adaptif Oleh Sistem Fuzzy. *Jurnal Masyarakat Informatika Indonesia (JMII)*, vol. 1, no. 1, pp. 25 - 30.
- F. Westad, M. Hersleth, P. Lea, H. Martens. 2003. Variable Selection in PCA in sensory descriptive and consumer data. *Food Quality and Preference Journal*, vol. 14 issue 5 – 6, pp. 463 – 472.
- G. Kaminsky, S. Lizondo dan C. Reinhart. 2008. Leading Indicator of Currency Crisis. *Munich Personal RePEc Archive No. 6981*.
- J. Han, Micheline Kamber, & Jian Pei. 2012. Data Mining Concepts & Techniques Third Edition. *Morgan Kaufman, Elsevier*.
- I. H. Witten, Eibe Frank, & Mark A Hall. 2011. Data Mining : Practical Machine Learning Tools and Techniques Third Edition. *Morgan Kaufman, Elsevier*.
- M. L. Brown & John F Kross. 2003. Data Mining and the Impact of Missing Data. *Industrial Management & Data System*, pp 611 – 621.
- S. Razzaqi et al. 2011. Dynamic Relationship Between Energy And Economic Evidence For D8 Country. *Pakistan Development Review*, vol. 50, pp. 437 - 458.
- S. Singhal & Monika Jena. 2013. A Study On Weka Tool For Data Preprocessing , Classification, And Clustering. *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, vol. 2 issue-6, pp. 250 – 253.