Identifying Differentially Expressed Genes in Different Stages of Lung Cancer – an application of ARM Model on Gene Expression Data

Authored By :

- Subir Hazra, Rohan Sarkar, Amartya Roy, Anupam Ghosh.

Presented By :

- Amartya Roy.

Overview

- Cancer & Gene’s.
- Molecular Biology and Microarray gene technology.
- Plan of execution.
- Related Works.

Association Rule

What is Association Rule?

Application of Association Rule (Market Basket analysis)

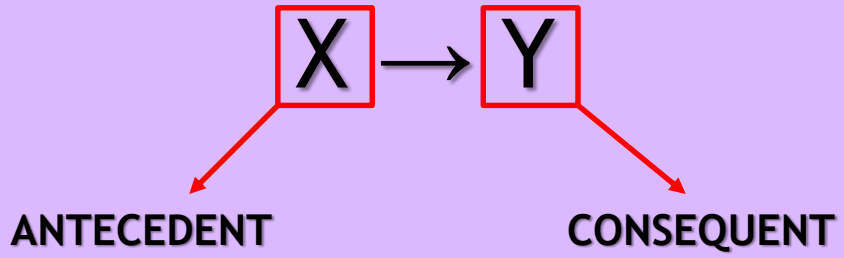
How we apply this Association Rule?

Terminology

Association Rules

An association rule is an implication expression of the form X→Y where X and Y are disjoint item sets

How Association Rule applied here?



Matrix Structure

Data Representation and matrix structure

	S1	S2	S3	S4
g1	0.68	-0.5	0.78	-0.34
g2	-0.23	1.2	0.61	0.89
g3	0.66	0.84	0.99	-0.10
g4	0.87	-1.0	-0.67	-0.44
g5	-1.0	0.83	0.65	0.61
...				

TID	items
T1	I1, I2 , I5
T2	I2,I4
T3	I2,I3
T4	I1,I2,I4

Sample Dataset

	Unnamed: 0	Unnamed: 1	Sample ID	AD10	AD2	AD3	AD5	AD6
0	NaN	NaN	Cluster ID	49.0	64.0	63.0	60.0	30.0
1	GENE	PROBESET	IN 4966 genes	NaN	NaN	NaN	NaN	NaN
2	GABRA3	A28102_at	YES	170.0	59.7	80.0	92.4	104.0
3	OMD	AB000114_at	YES	69.4	18.1	26.0	96.9	72.8
4	GS3686	AB000115_at	YES	250.7	146.8	150.0	177.8	228.7

7129X96

	AD10	AD2	AD3	AD5	AD6
2	170.0	59.7	80.0	92.4	104.0
3	69.4	18.1	26.0	96.9	72.8
4	250.7	146.8	150.0	177.8	228.7
5	957.1	186.8	340.2	515.8	540.8
6	25.4	-7.7	-16.3	18.0	26.0

7129X86

	LN64	LN66	LN67	LN69	LN70	LN71
2	108.0	108.0	170.8	88.0	97.0	89.1
3	93.0	103.5	100.5	72.8	118.0	118.0
4	152.8	161.4	292.1	109.8	215.8	325.9
5	387.3	347.2	702.8	643.8	368.9	429.4
6	32.0	3.9	-1.7	19.3	39.0	-4.0

7129X10

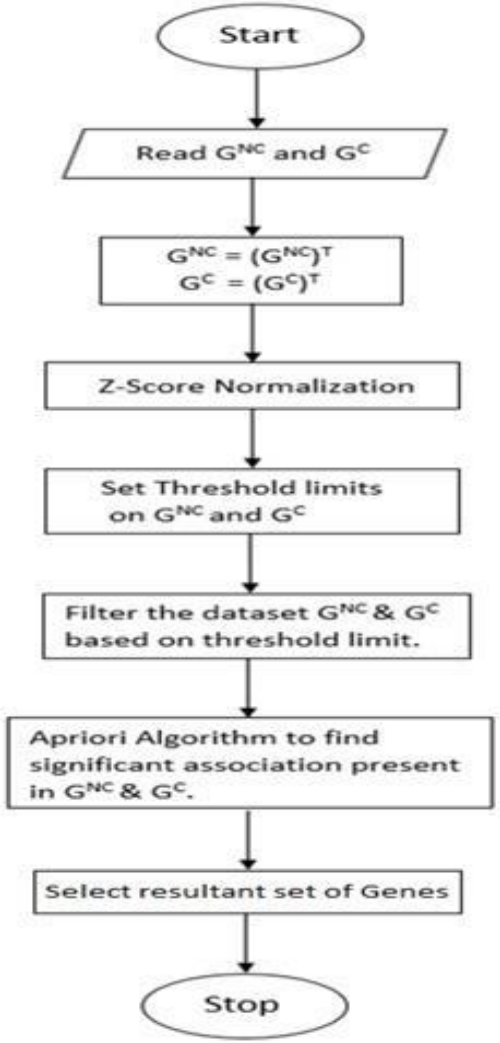
	AD10	AD2	AD3	AD5
2	170.0	59.7	80.0	92.4
3	69.4	18.1	26.0	96.9
4	250.7	146.8	150.0	177.8

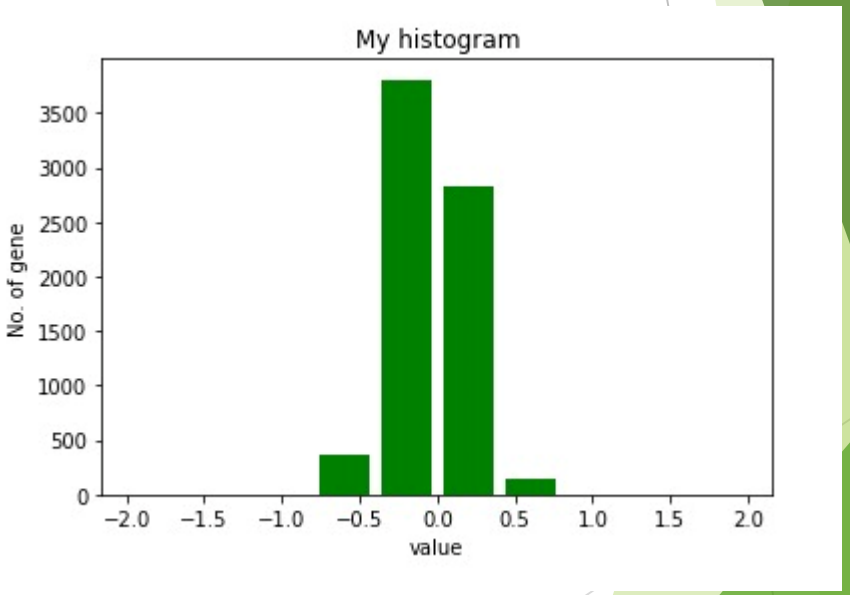
7129X67

	AD10	AD2	AD3	AD5
2	170.0	59.7	80.0	92.4
3	69.4	18.1	26.0	96.9
4	250.7	146.8	150.0	177.8

7129X19

Methodology

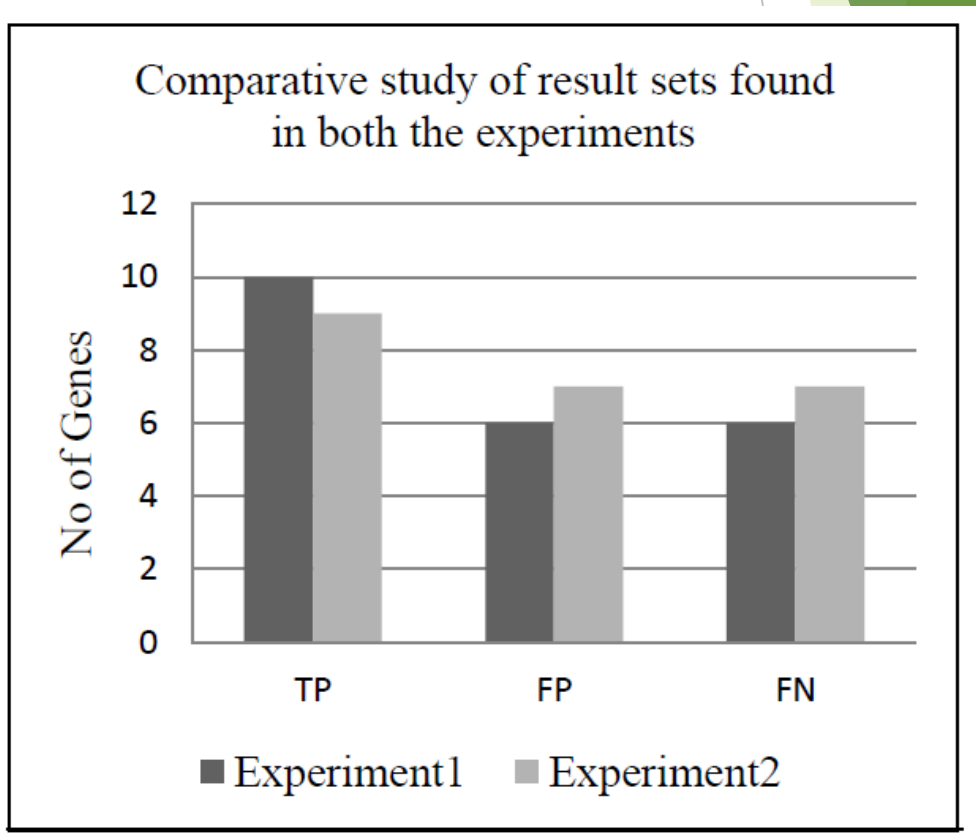




Results

- We find Association's present in Normal state but not in Cancerous state.
- Newly formed Association in Stage I & Stage III.
- Significant Gene in both Stage I & Stage III.

Significant genes which are differentially expressed	
Normal state to cancerous state(stage-I)	*Stage-I to stage-III*
MIG	ASGR2
TNFAIP2	APM
S100A2	PRSS1
AMP1	SERPIND1
CHI3L1	SERPINB5



Conclusion

Systematic and unbiased approach to cancer classification is of great importance to cancer treatment and drug discovery.

In our paper, we have reviewed Apriori algorithm for mining frequent pattern from microarray gene expression data

Recent studies have shown that gene expression changes are related to different types of cancer

References

- Hu W, Li X, Wang T, Zheng S. Association mining of mutated cancer genes in different clinical stages across 11 cancer types. Oncotarget ;7(42):68270-68277, 2016.
- KALE, O. V. and MOMIN, B. F. "ASSOCIATION RULE MINING FOR GENE EXPRESSION DATA," International Journal of Computer Science and Informatics: Vol. 3 : Iss. 4 , Article 4, 2014.
- Chen, SC., Tsai, TH., Chung, CH. et al. Dynamic association rules for gene expression data analysis. BMC Genomics 16(1), 2015.
- Singh, Prithvi & Dohare, Ravins & ali, Sher et al. Identification of differentially expressed genes in small and non-small celllung cancer based on meta-analysis of mRNA. Heliyon. 5(6), 2019.
- Maran, Pyingkodi, Thangarajan R..(2018). Informative Gene Selection for Cancer Classification with Microarray Data Using a Metaheuristic Framework, Asian Pacific Journal of Cancer Prevention : APJCP, 19,561-564, 2018.
- Zihayat, Morteza & Davoudi, Heidar & An, Aijun. Mining significant high utility gene regulation sequential patterns. BMC Systems Biology. 11(S6), 2017.
- Al-Aamri, Amira & Taha, Kamal & Al-Hammadi, Yousof et al. Analyzing a co-occurrence gene-interaction network to identify disease-gene association. BMC Bioinformatics. 20(1), 2019.

Thank you