7-4-	0	31 3	Dulle 1	AHAND	
1	00	N.	0 00		~ ()
W		M		PU	5 3
V	1001		0		
72-	und all		tort	2019th	of the

- PROBLEM STATEMENT: Apply a-priori algorithm to find frequently occurring items from given data and generate strong association rules using support and confidence thresholds.
- objective: model associations between products by determining sets of items frequently purchased together, and building association rules to derive recommendations
- outcomes: Demonstrated market basket analysis using a-priori algorithm to find frequently occurring items from given data and generate strong association rules using support and confidence thresholds.
- processor or higher), 4GB RAM or higher, 1GB HDD or 1280BSSD minimum; 64 bit UNIX (LINUX OS, Pythors, jupyter, apyoni

DD THEORY

Association Rule Mining finds interesting association and relationships between large sets of data items. This rules showings how frequently an itemset occurs in a transaction. It is defined as an implication expression of the form $X \rightarrow Y$, where X, Y are any 2 itemsets.

Market Basket Analysis is one of the key techniques used by large relations to show associations between items. It allows retailers to identify relationships between the items that people buy to together frequently, so that items can be strategically placed next to each other to boost sales.

Some definitions:

4 Support count: frequency of itemset occurrence to Frequent itemset: itemset whose support >=

The Association Rule Evaluation Metrics are:

- 1) support (3): The number of transactions that include items in the EX3 and EY3 parts of the rule as a percentage of the total number of transaction.
- 2) confidence (c): It is the ratio of the number of transactions that includes all items in an itemset to the as well as the transactions that include all items in {A} to the number of transactions that includes all items in {A}.
- 3) Lift (1): The lift of the rule $x \rightarrow Y$ is the confidence of the rule divided by the expected confidence, assuming the confidence such that itemsels $\{x\}, \{Y\}$ are independent of each other. The expected confidence is the confidence divided by the frequency of $\{Y\}$ for Itemsels occur together: i) as expected if lift = 1,

	ii) more than expected if lift >1, and iii) less than expected if lift <1.
0 21	APRIORIA ALGORITHMINE LITTORIAN SUIT
	Apriori is an algorithm for frequent item set mining and association rule learning over relational databases, first proposed by Agrawal and Srikant in 1994.
	It is designed to operate on databases containing transactions, with each transaction being a set of items (itemsets).
	Given a threshold C, the algorithm identifies the item sets which are subsets of at least C transaction -s in the database.
	It uses a 'bottom up' approach, where frequent subsets are extended one at a time (called candidate generation), and groups of candidates are tested against the data. The algorithm terminates when no other successful extensions are found.
	A-priori uses breadth first search and a hash tree to count candidate item sets efficiently. It generates item sets of length k-1; then it prunes the candidates which have infrequent sub pattern. After that, it scans the transaction database to determine frequent item sets among the candidates.

	male than expedied if the 21 and ii) ress th
	CONCLUSION:
	SONCEOGION
	Thus frequently occurring items from given Market
	Basket dataset and strong association rules using
102 1	support and confidence thresholds we're found using
	a-priori algorithm.
	databases first proposed by Agraigal and St
View land	1904.
A	
	It is designed to operate an dolabases?
+	minsactions, with each transaction being a sc
	items (itemsets).
- JR - 500	thank mathonia at a bladzinat a ciaum than than the
	Them sets nothing are subsets of all least a
	e in the database.
imain	. It mees a bottom was approach cohere he
	subsets are exercted at a tire confer
310 E 311	with the for the first the first state of the first
	rested against the data. The algorithm trans
	when successful extensions one found
	A -pajori uses breadth first search and a h
	HOSTOFF 2-128 GARAGORE HORO CIT 29T
2-2011/0	The government of headth k-1; then a
	the condidates usual house introquent out to
- 11	After that it soons it control that control and a delevation from the sound of the
-31001	determine frequent man, sets among the con-