SOCIAL NETWORK MINING

Nishanth.R.Lawrence(1060910054) &
Vivek Kumar Ranjan(1060910411O)

ABSTRACT

To develop an efficient way for mining social networks by relational models i.e graph based nosql , text mining and crawling .

OBJECTIVE

Exploring oceans of data and building simple explanations for what we observe will lead to exciting new insights and discoveries that can simplify people's lives and decisions.

EXISTING SYSTEM

- Traditionally Social network mining has dealt with the data which are not networked.
- The dimensional(increasing size) problem is of critical scenario for analysis of social networks
- Dealing in aggregation in tables with terabytes of data prohibiting ,if the analysis has to useful for fast decision

EXISTING SYSTEM(continued….)

- Data Mining in Social Networks by David Jensen and Jennifer Neville
- Web Mining and Social Networking techniques and applications by Guandong Xu , Yanchun Zhang , Lin Li
- Mining Social Networks : Uncovering Interaction Patterns in Business Process by Minseok Song

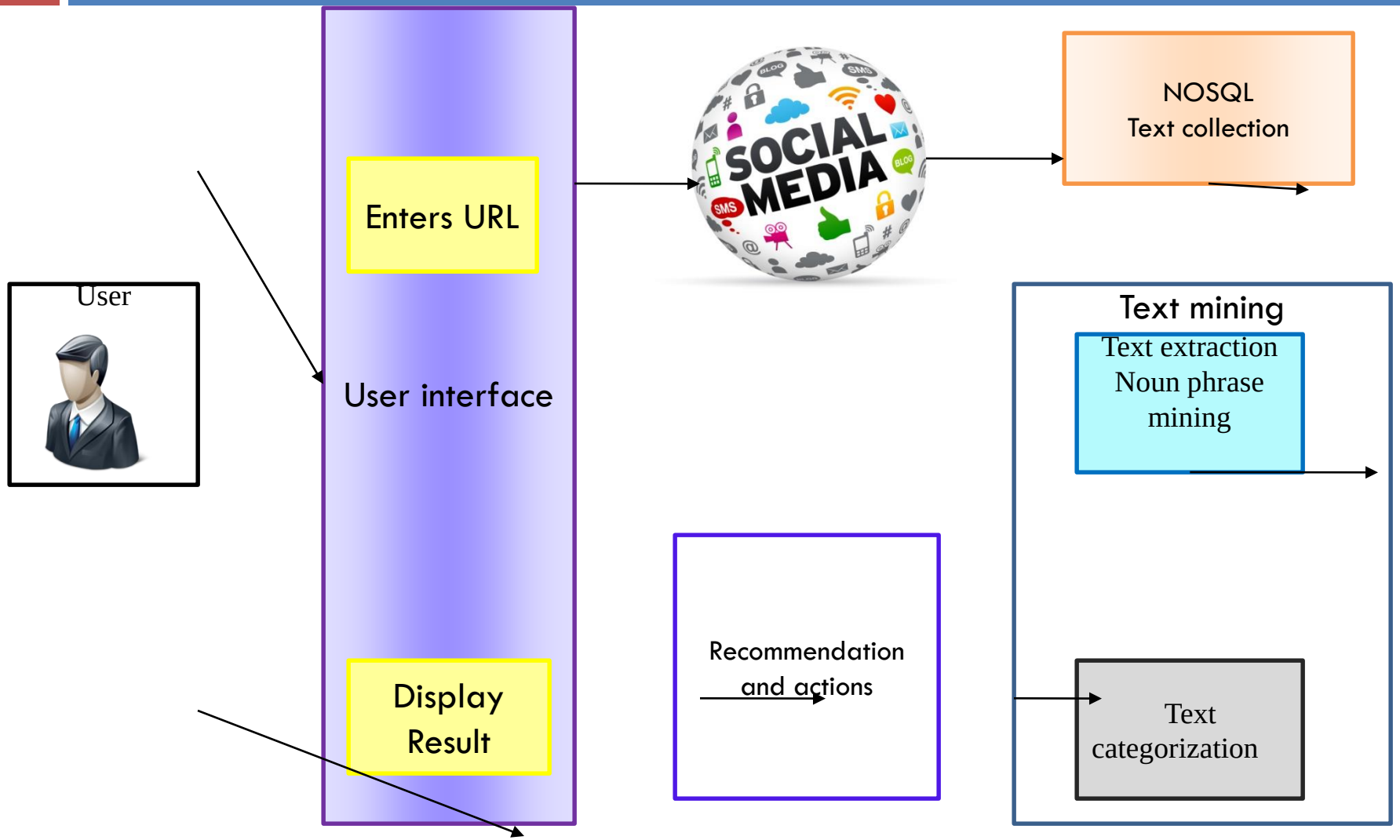
PROPOSED SYSTEM

- Scalable (Can handle any size)
- Overcomes known sql limitations
- Easy to adopt
- Schema free

PROPOSED SYSTEM(continued….)

- predicting poll results depending on the views of the people
- predicting the reach among the public for an ad or a product .
- Avoiding juvenile delinquencies by finding out the kids who have developed hatred towards school or community etc
- Avoid cyber crime by monitoring abuse words .

ARCHITECTURAL DESIGN



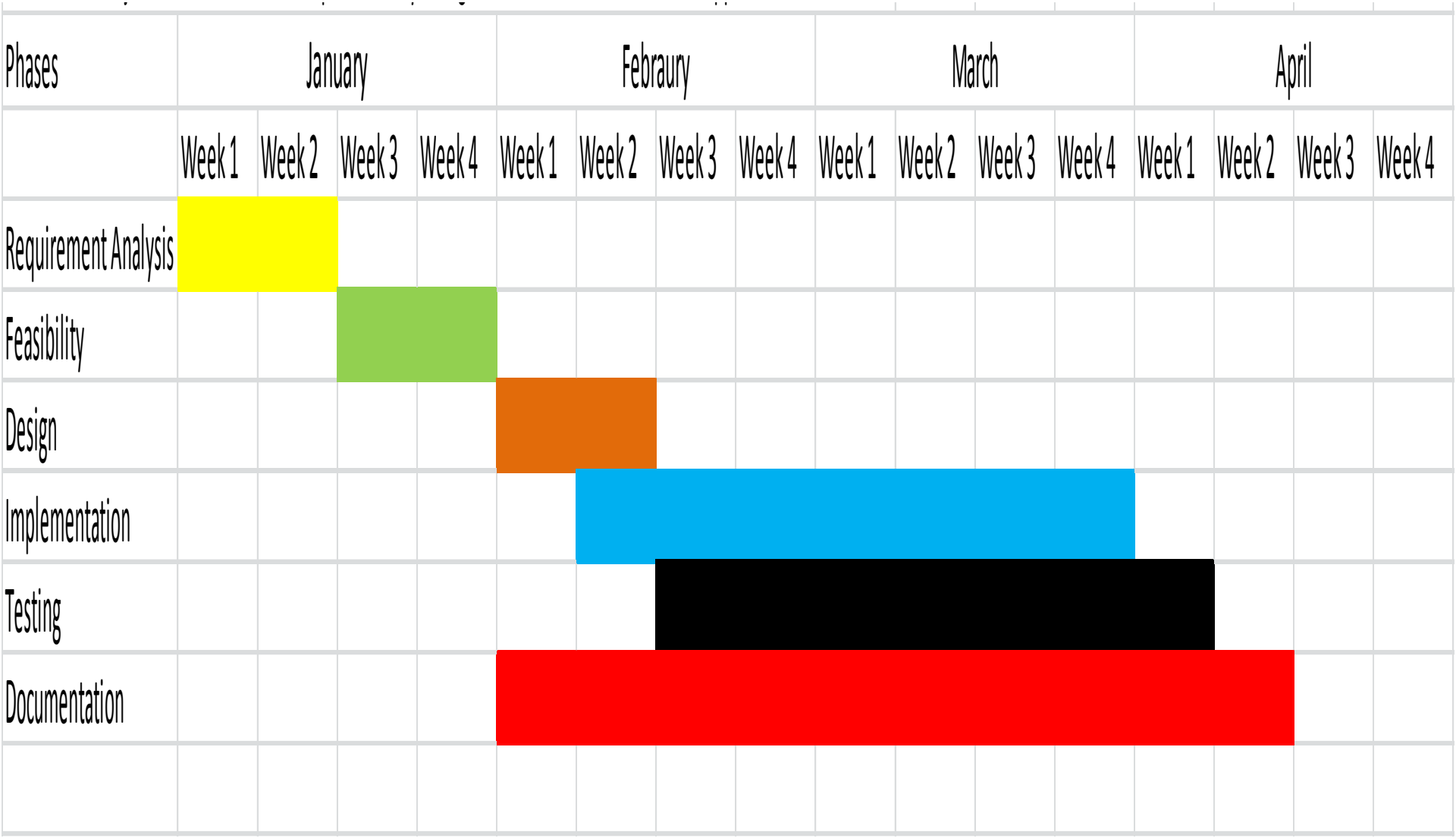
LITERATURE SURVEY

- Due to its graph data model , is highly agile and blazing fast
- For connected data operations, the use of graph based nosql makes it run a thousand times faster than relational databases.

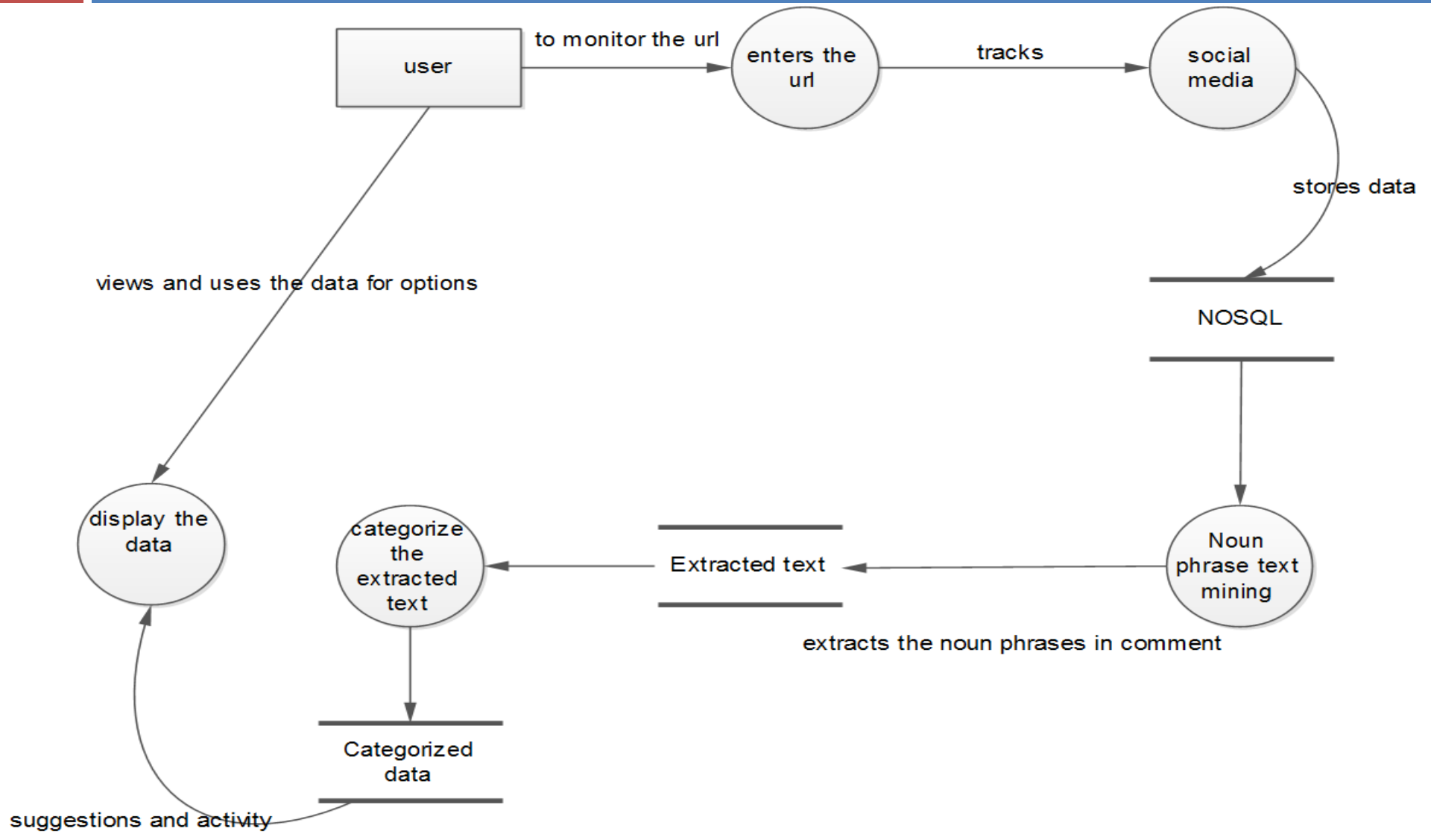
MODULE SPLIT-UP

1. INPUT : Nosql storage
2. (web crawling)
3. PROCESSING :Pre-processing of data set
4. (Various data mining process)
5. OUTPUT :Displaying result in a graph model

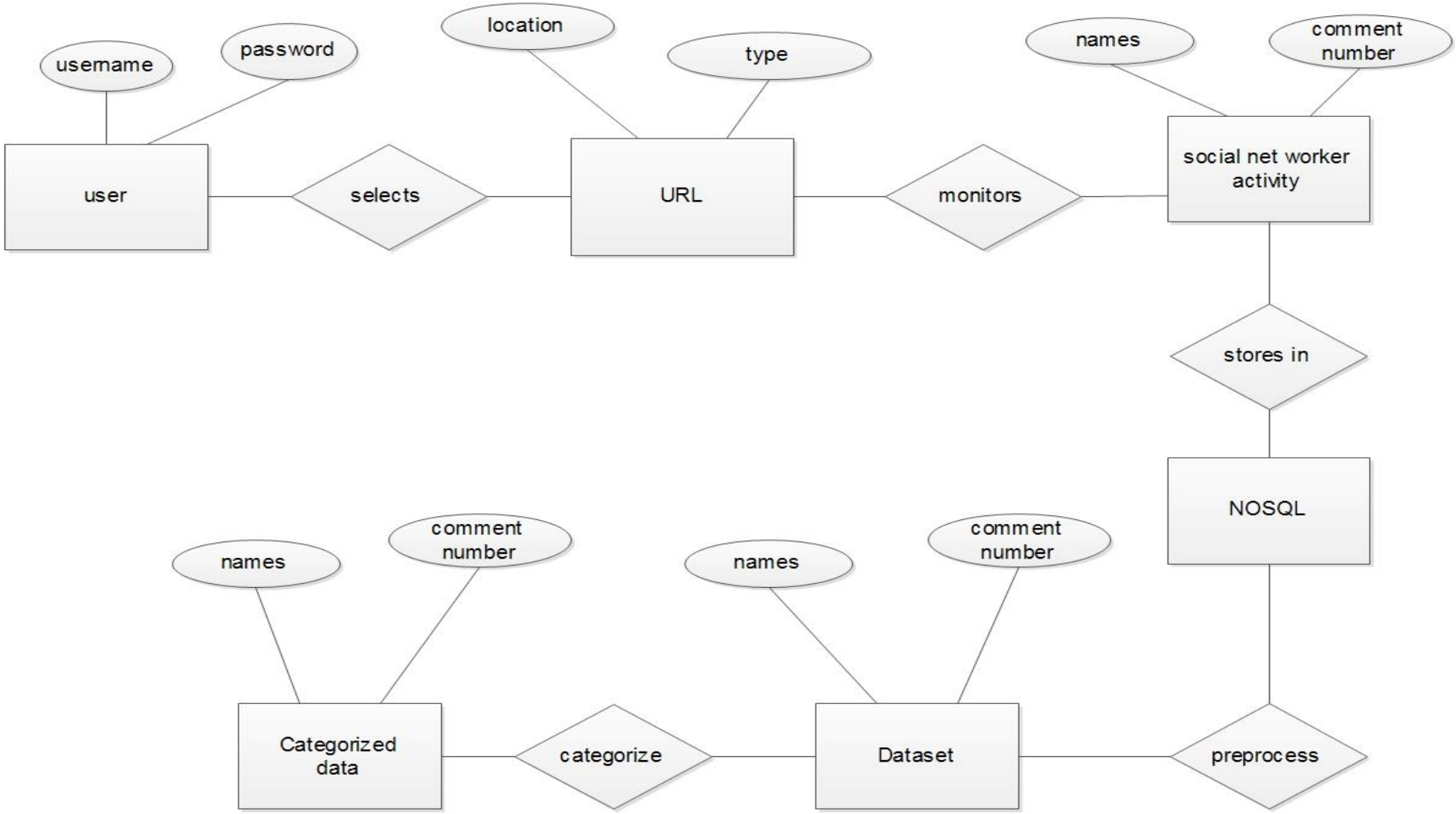
GANTT CHART


Phases
January
Febraury
March
April
Week 1
Week 2
Week 3
Week 4
Week 1
Week 2
Week 3
Week 4
Week 1
Week 2
Week 3
Week 4
Week 1
Week 2
Week 3
Week 4
Requirement Analysis
Feasibility
Design
Implementation
Testing
Documentation

DATA FLOW DIAGRAM



ENTITY RELATIONSHIP DIAGRAM



USE CASE DIAGRAM

ALGORITHM USED WITH COMPLEXITY

- The k-means algorithm
- Support vector machines
- The Apriori algorithm

EXPECTED OUTCOMES

- A model showing relationship between different objects

REFERENCES

□[1] P. Domingos and M. Richardson. Mining the network value of customers. In Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining,pages 57{66, San Francisco, CA, 2001. ACM Press.

□[2] R. Dye. The buzz on buzz. Harvard Business Review, 78(6):139{146, 2000.

□[3] S. Jurvetson. What exactly is viral marketing? Red Herring, 78:110{112, 2000.

□[4] D. Kempe, J. Kleinberg, and E. Tardos. Maximizing the spread of inuence in a social network.

□In Proceedings of the Ninth ACM SIGKDD

THANK YOU