

**PROJECT REPORT**

**Twitter Search Application**

**Prepared by**

Vishnuram Jatin Bangaru(vb440)

Ankeeta Priyam(ap2213)

Noopur()

Ganesh Raj()

Table of Contents

[**INTRODUCTION** 3](#_Toc131969136)

[**PERSISTED DATA MODEL AND DATASTORES** 5](#_Toc131969137)

[**PROCESSING TWEETS FOR STORING IN DATASTORES** 6](#_Toc131969138)

[**SEARCH APPLICATION DESIGN** 7](#_Toc131969139)

[**RESULTS** 8](#_Toc131969140)

[**REFERENCES** 10](#_Toc131969141)

# **INTRODUCTION**

**DATASET**

# **PERSISTED DATA MODEL AND DATASTORES**

Describe your user and tweet data model in detail (preferably with diagrams).

Describe how you stored the data in the datastores.

What did you optimize for? What were the tradeoffs?

Did you create any indexes for the datastores?

# **PROCESSING TWEETS FOR STORING IN DATASTORES**

Make sure that you processed the tweets and stored the information (users, tweets) in the datastores one at a time (DO NOT load all the data into a dataframe and then load the dataframe into the datastore)

# **SEARCH APPLICATION DESIGN**

What types of searches are allowed? What kinds of drill-downs are allowed?

How were search queries translated into queries for the datastores?

What was your notion of relevance (i.e. how did you order the results of the search queries)?

How is the cache used?

# **RESULTS**

Show the results of each type of query.

Timings of your test search queries (make sure you are hitting cached and non cached data)

**CONCLUSIONS**

# **REFERENCES**