MOM - 28 Feb 2020

Problem statement -

**which reciepe belongs to which region? (Based on Intgradients)

**literatue reviw based on context vector concept atleast (8 papers in text analytics)

** If possible find the datasets

Share chosen papers by Monday 2/3/2020 EOD

1 paper should concentrate on theorotical aspect and 1 on technical

Summerize these paper.

Nishant to share the excel sheet

Next Meet - Important Stuffs - 5th March 12:00

Abstract, Introduction, Expermintation, Evaluation steps, We can add results after 9th March

MoM 5th March

Ideas Discussion - Need to improvise so guys try to break it from all possible points

We have dish A

It has ingredients [1,2,3]

Recipe A has 1000 reviews

So we do text analytics on those 1000 reviews and generate binary label for that dish A
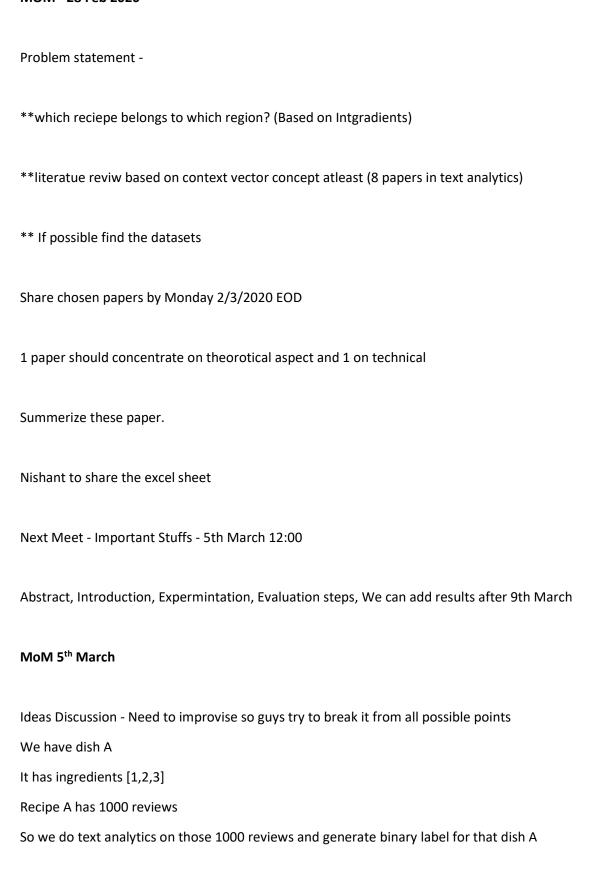
Positive

Negative

So now we have

Data row like

1. Dish A

[1,2,3] = 1

Assume 1 positive

So we have n dishes represented by context vectors of ingredients and their labels

So now when we get a new dish B

And it's ingredients

[1,3,4] we can predict if these

Dish or combination of ingredients would be linked or disliked by the customers/ppl

MoM 7th March 2020

Research questions

A list of research questions we would like to address during the project:

What are the most commomly used ingredients?

What are the most distinctive ingredients?

How can we represent cuisines as a network of ingredients ?

How can we cluster cuisine in terms of their recipe components?

How do cuisines influence each other?

MoM 15 March 2020

Scope of the project -

Can we predict the contents on the quizine based on reciepe/intgradients context vectors?

Lujain - Continiue to build taxanomy

Aakash - Word2Vec check for the cosine distances (Wikipedia Taxonomy)

Nishant - Data cleaning

Oommen – Understand Word2Vec and Glove Methodlogy

Rohan - To create Context vectors

MoM 25 March –

Covid – Continue working on the same.

Nishant Plus Rohan – Check the feasibility of the code.

MoM 2nd April 2020

Lujian --→ Nishant/Rohan -→ Ommen/Aakash

Lujian to generate data…Send data to Nishant to clean and rohan to build vectors, Ommen/Aakash to Generate Charts and visulizations

Next Meet – 6 April 5pm

MoM 6th April 2020

Drop an email - Modification in the problem statement(discard taxonomy)

Compare results based on Word2Vec and GloVe - Nishant and Rohan (Implementation and Evaluation)

Look into the Visulization Task (Reciepe Vectors) - Oommen

Find existing paper on Bag of Words. Explain in the paper as well

Modification in Literature Survey (Luijan and Aakash)

     - (Try to answer 15 Ques in the list)

     - Modification in Abstract, Introduction, keywords and Formatting

     - Previous work on frequency based models, word2vec, etc, etc

Psuedo Code - Word2Vec, GloVe and Our Algorithm

MoM 11 April –

Lujain, Aakash - Continue Building the paper

Nishant, Rohan - Test Context Vectors

Oommen – Continue building cluster visualization

Next Meet 13 April 13, 2020

MoM 2nd April 2020

Proofreading and Final Actions

Submit Articles with solving below errors -

1. Change in title!

From: Classifying Recipes into Cuisines Using Context Vectors

To: Classifying Recipes into Cuisines Using Context Vectors of Ingredients

2. Abstract first line the word "like" -> "such as"

3. Line no 5 Abstract: of cuisines using the context vectors -> of ingredients

4. In abstract which baseline method: traditional frequentist approach using bag of words

5. Intro second para last line: a small corpus per entry

6. Intro third para " We will be following that with our results, analyse them along with any errors. "

7. e large corpus, like Wikipedia section 2.1.1

8. 2.2 related work no space

oss the world have attempted to predict a dish's cuisine based on a recipe's ingredients as

9. related work : xgnoost mai dala but what was put words or count vectors

10. Sectoi 3.2

. The prediction accuracy for this method was found to be 78%. (classification)

11. 3.3 Cleaning Up The Data

we first remove

12. by adding should we mention

13. 3.4

Not split randomly