Business Analytics Centre Symposium (BACS) 2020
26 September 2020 | Singapore

Automated Financial Commentary Generation

Presented by: Utkarsh Chaturvedi

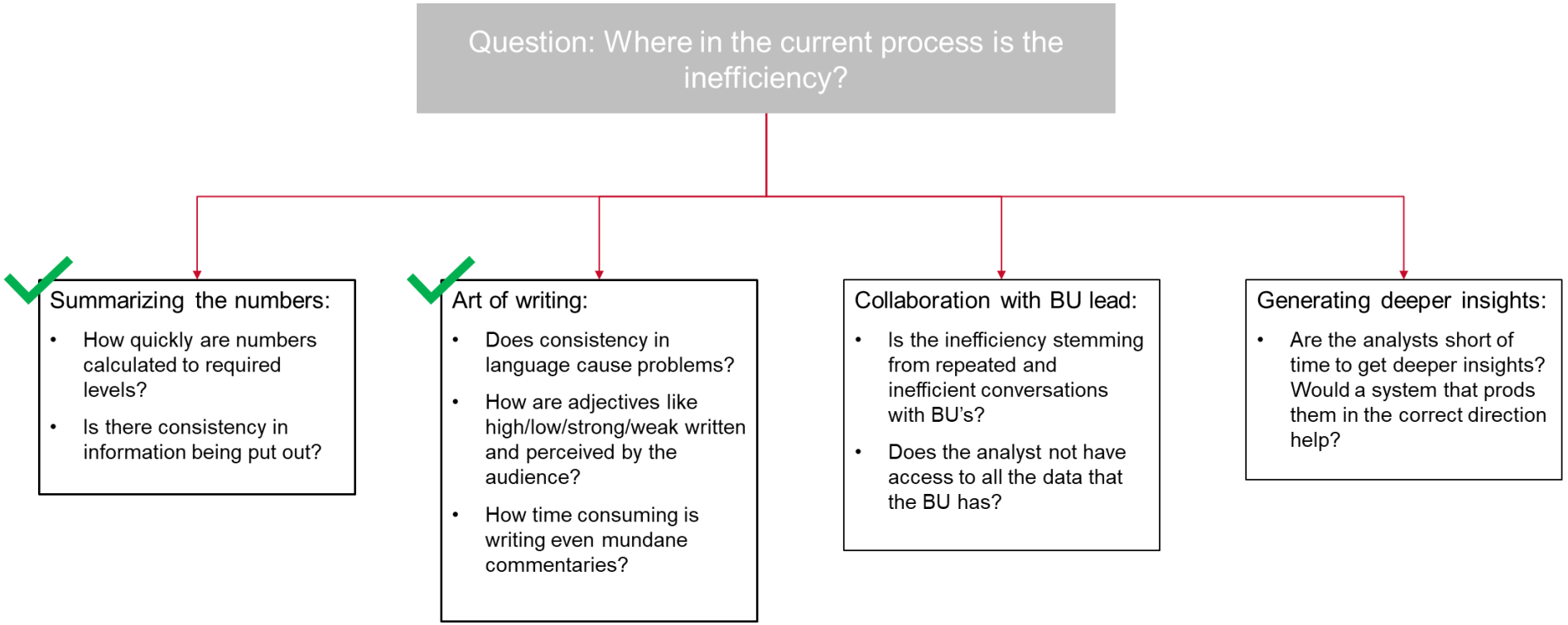
Supervised by: Prof. Teo Thompson

NUS National University of Singapore

NUS Business Analytics Centre
SUPPORTED BY NUS BUSINESS SCHOOL & NUS COMPUTING

What was the business problem?

- To help financial analysts **increase their efficiency** by **automating the task of writing lower level commentaries** and re-directing their efforts on identifying causality of important financial numbers



What was the solution approach?

The process was broken down to 4 separate components:

1. Analysis of the language
2. Building codes to ensure flexibility of usage across product hierarchies, geographies and BU's
3. Using NLG tools to realise the language
4. A front-end to facilitate ease-of-use of product

Deep-dive into the tool

Step 1: Analysis of language

Basic → Advanced

Word Cloud

- Overview of the most commonly used words
- Overview of words as we move across different dimensions of data i.e. month/plan/period

Language Analysis (using NLTK)

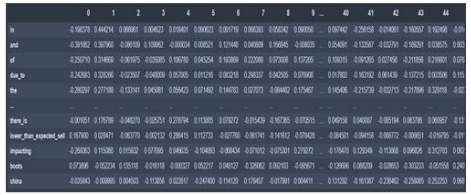
- PoS tagging to further understand the language being used
- Identify syntactically similar words

Topic Modelling

- Attempt to split comments based on their complexity and depth
- Topics appeared to be 'mathematically' different rather than based on human interpretation

Word2vec model

- Tries to learn the meaning and sense of the word using context of words around it
- Identify words implying similar things



Deep-dive into the tool

Step 2: A technical coding challenge

- The solution was built with a key requirement of being flexible and robust to variations in data input structure, product and geographical hierarchies across business-units and ease-of-use to analysts
  - As a result, there could be no-hardcoded columns or values within the entire code

Deep-dive into the tool

Step 3(i): Generating the language

Using Templates

- A pre-defined template has been set up where relevant information is simply plugged in
- With the diversity of analysts, BU's and regions, a template fit could seem too mechanical

Non-neural NLG

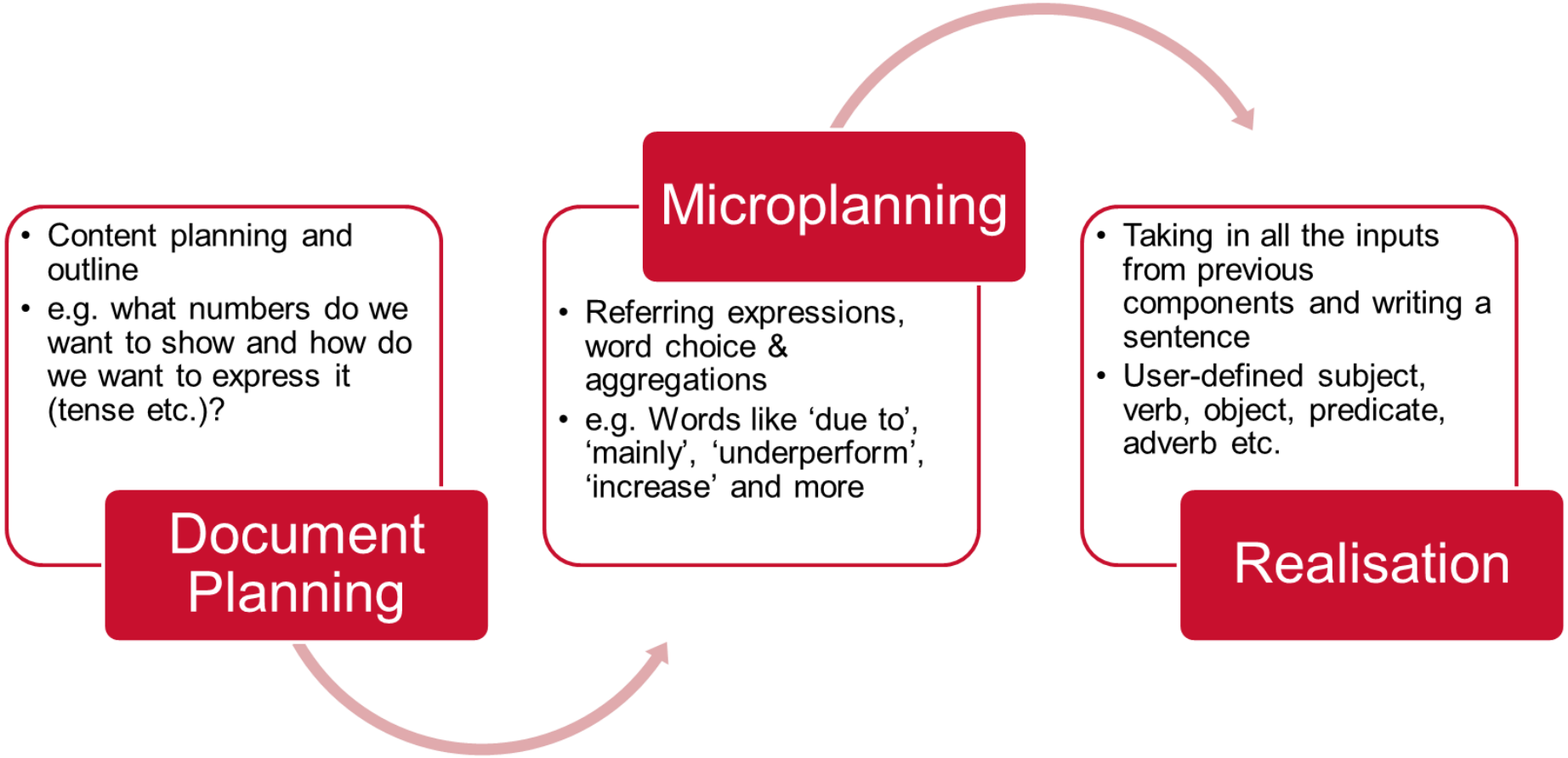
- An older but reliable method to generate natural language
- Breaks the process down to Document Planning, Microplanning & Realisation
- Current commercial tools in market are built on this

Neural Network based NLG

- End-to-end Data-to-text NLG solutions are at the cutting-edge of research, with developments in few-shot and zero-shot networks
- Lack of convex convergence can make this unreliable
- Does not have any extensive commercial applications yet

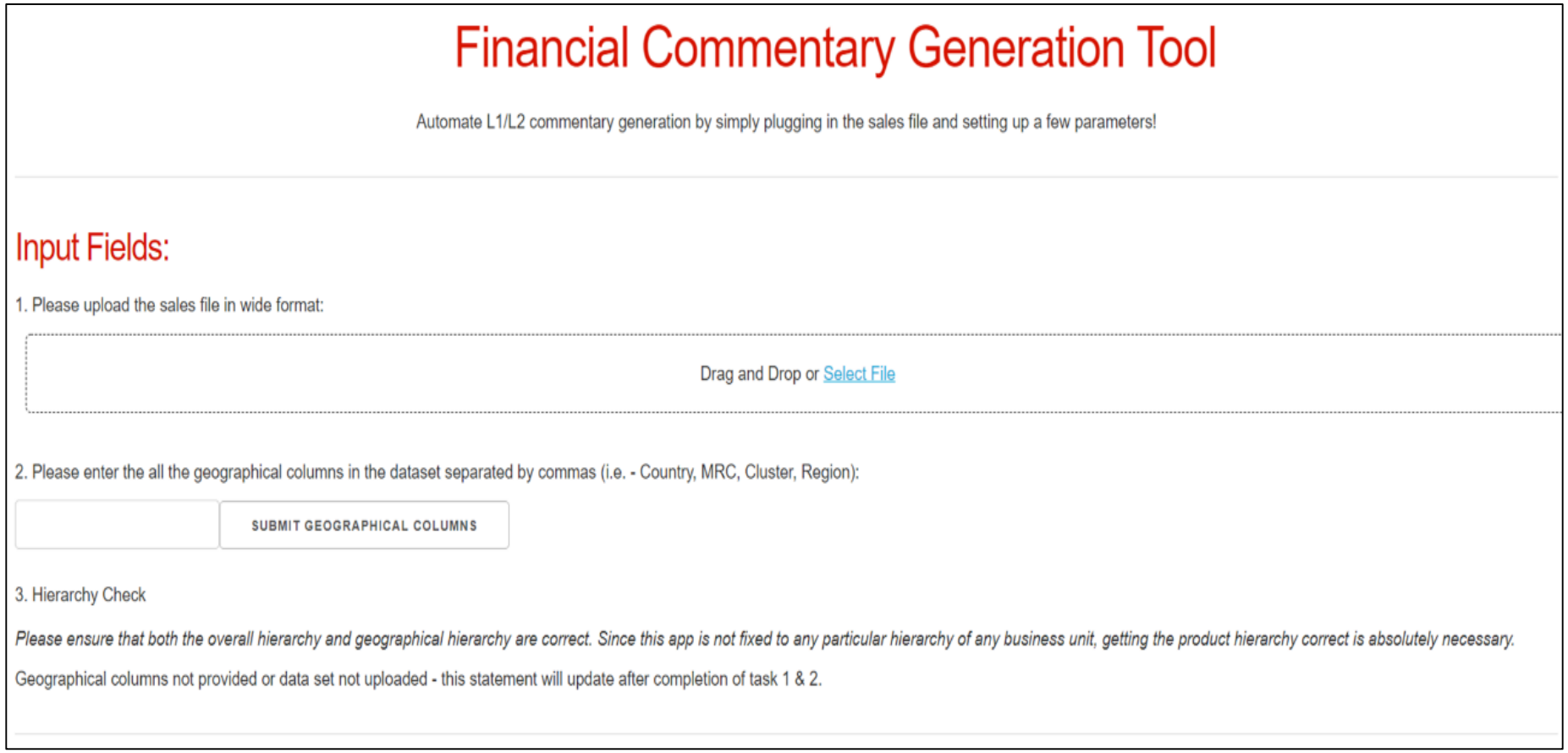
Deep-dive into the tool

Step 3(ii): Non-neural NLG - how does it work?



Deep-dive into the tool

Step 4: Dash-plotly based web-tool



Deep-dive into the tool

Features of the tool:

1. Flexibility to user
   a) No hard-coded features or columns* - user can use any columns names/hierarchies provided data some bare necessary columns such as NTS/Period/Month are present
2. Easy to modify front-end for different needs across BU's/geographies
3. Modular function-based backend allowing for easy modifications in further iterations
   a) The current codes have been written to ensure that further feature addition takes minimal effort

What next?

1. Showcase current capabilities and try to incorporate feedback from the end-user(s)

    a) Set up meetings with multiple stakeholders to raise awareness and get their buy-in

2. Diversify data sources and work towards generating advanced level commentaries

    a) Need to have further discussions with the analyst on the sources of information and how it is processed to arrive at the comment

Utkarsh Chaturvedi
+65-84117760
utkarsh.chaturvedi@u.nus.edu
utkarshchaturvedi95

THANK YOU!

NUS National University of Singapore

msba@nus.edu.sg

msba.nus.edu.sg

@nusmsba

nus msba

NUS Business Analytics Centre

SUPPORTED BY NUS BUSINESS SCHOOL & NUS COMPUTING