

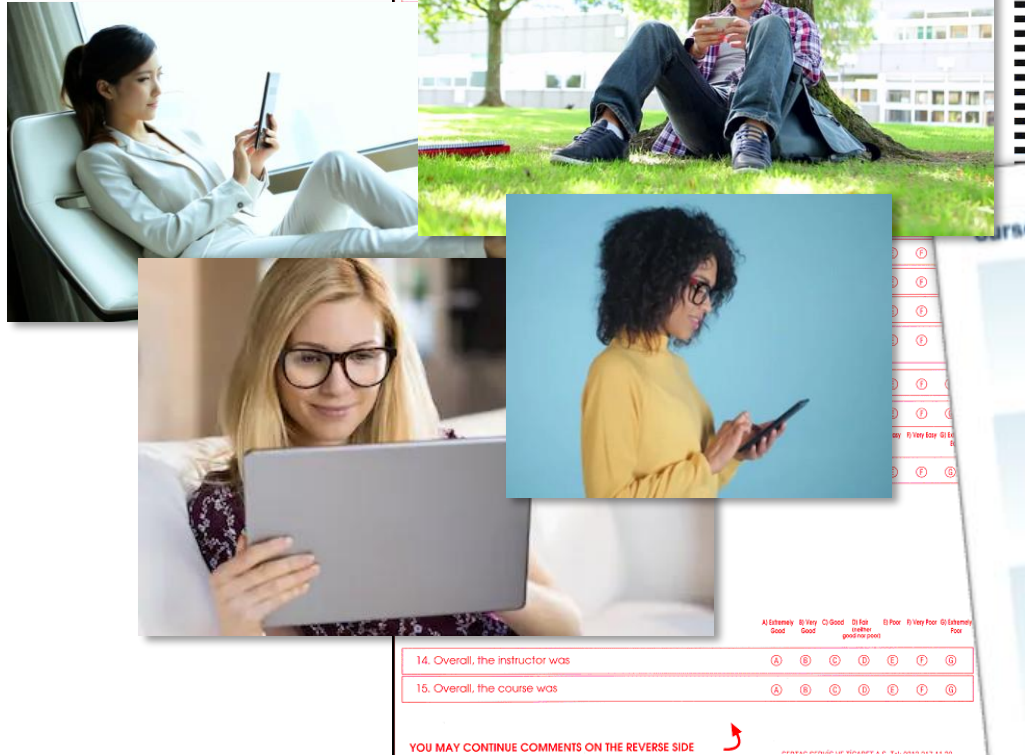
Unveiling the Blue Text Analytics machine learning project

- *Presented by : Long Hua*

AGENDA

1. Overview of BTA concepts
2. Discussion of Natural Language Processing (NLP) applied to BTA

Shift in habits...



ABC UNIVERSITY **A 1.**
Course-Instructor Survey

COURSE UNIQUE NUMBER
 0 0 0 0 0 0
 1 1 1 1 1 1
 2 2 2 2 2 2
 3 3 3 3 3 3

INSTRUCTOR'S NAME AND SURNAME
 INSTRUCTIONS
 Enter your instructor's name and the course information in the box to the left and right.

14. Overall, the instructor was A B C D E F G

15. Overall, the course was A B C D E F G

YOU MAY CONTINUE COMMENTS ON THE REVERSE SIDE

Open-ended feedback

Digital Quantity
 Quality Contextual

Course Component Evaluation 2018 for 401791 African Cultures and Societies

The course was intellectually challenging

Strongly Disagree
 Disagree
 Neutral
 Agree
 Strongly Agree

The instructor was well-prepared for class

Strongly Disagree
 Disagree
 Neutral
 Agree
 Strongly Agree

How could the learning experience of this course be improved?

Listen the power points, or at least only use them as an aide. Focus on your lecture style. You describe the lecture good professor synthesizes the material, gives insight. It is frankly somewhat offensive that Dr. Thompson does not take more care in her lectures. The first two little paragraphs. This is high school level teaching.

Format lab 1 differently. I people doing different parts of the labstation doesn't make sense when we should all have a comprehensive understanding of the process, also sometimes the manual was confusing in terms of when to put what in which tube, like choose and put in a tube but it doesn't specify if it goes in the same one as the other stuff or if its two tubes.

BLUE TEXT ANALYTICS

BTA is

Efficient,
insightful
analysis on
large volumes
of open-ended
feedback



BTA Dictionary

A sampling of BTA

Themes and synonymous
expressions

APPROACHABLE:

approachable, calm, easy to talk to, laidback, good natured, is open to, open minded, relaxed

UNAPPROACHABLE:

aloof, chilly, detached, didn't care, distant, frigid, haughty, heartless, indifferent, intimidating, not easy to talk, overbearing, standoffish, stuck up, taciturn, too formal, unapproachable, unresponsive, unsympathetic

CLEAR:

accurate, clear, exact, explicit, precise, straightforward

UNCLEAR:

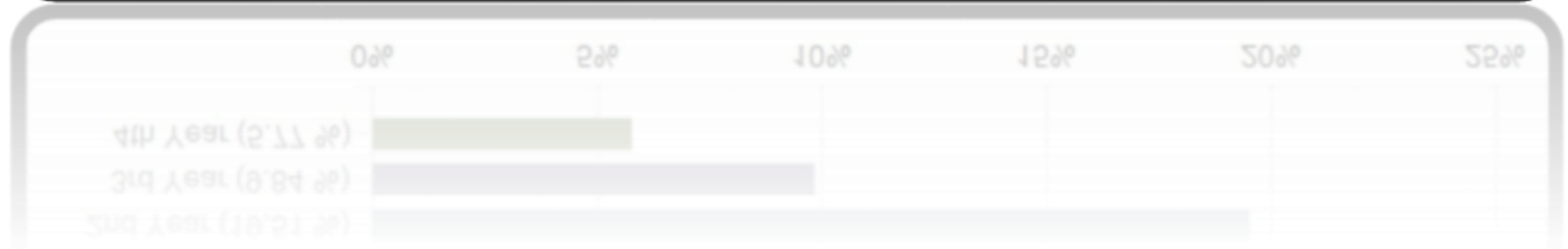
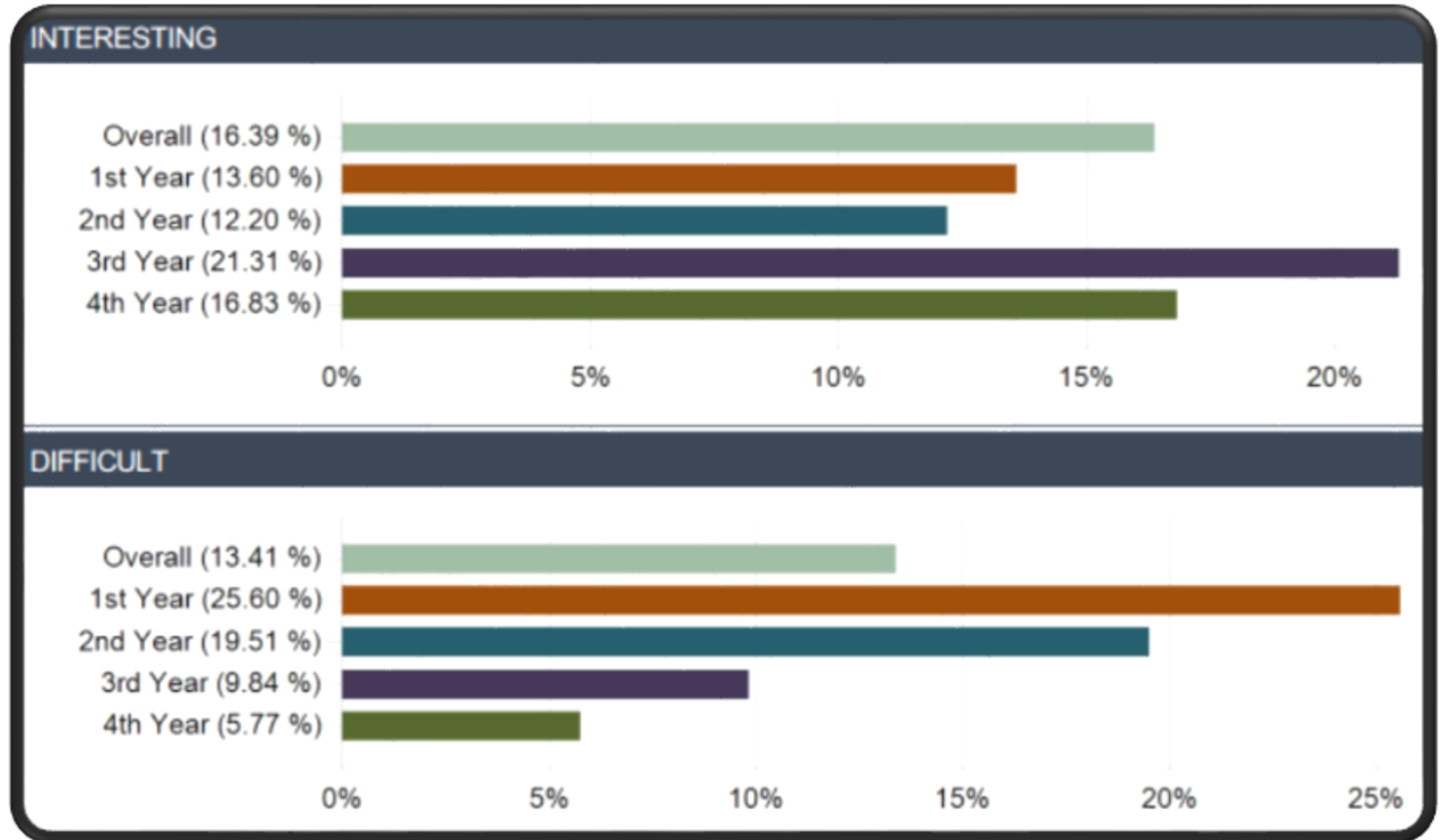
ambiguity, ambiguous, baffled, confusing, bewildered, blur, cryptic, enigmatic, vague, foggy, fuzzy, hazy, illegible, imprecise, in the dark, incomprehensible, inconsistent, lack of clarity, lost, muddy, not detailed, not explicit, obscure, opaque, puzzled, unclear, not well described

BTA Analysis

Themes compared across demographics

Comparing:

- a) Top 2 Attributes mentioned in the program
- b) Student year of study demographic data



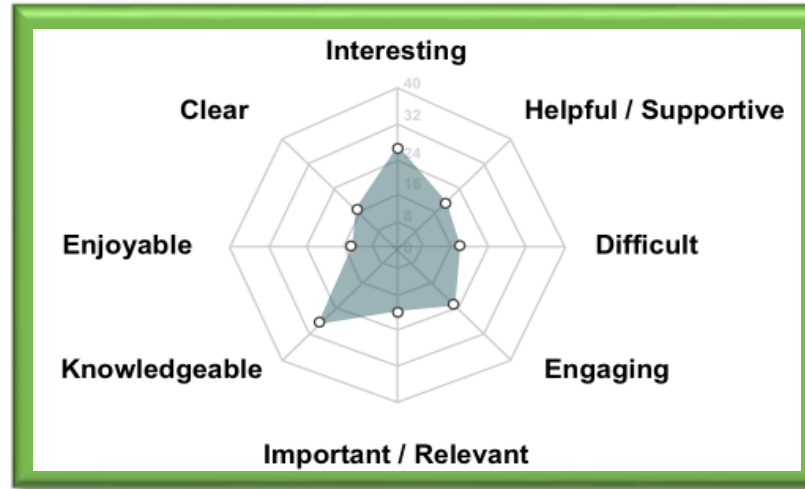
BTA Analysis

Themes compared across a rating question

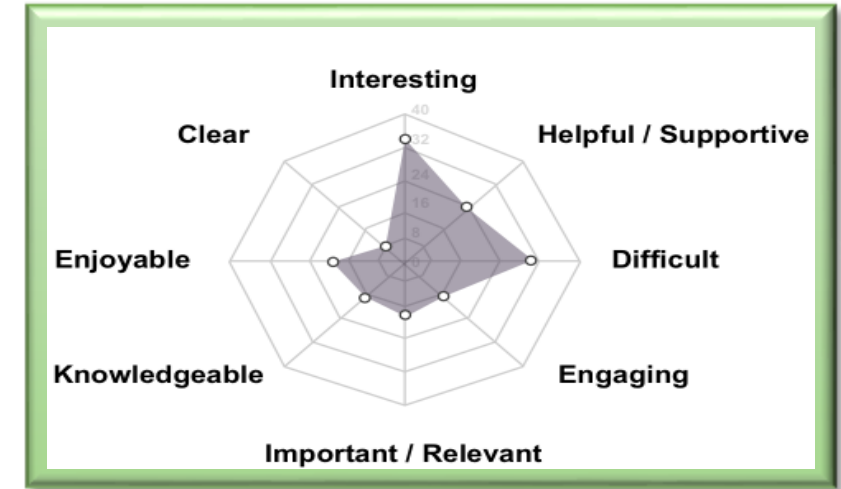
Comparing:

- a) Top 8 *Attributes mentioned* in the program
- b) "Overall what did you think of this course?" rating question

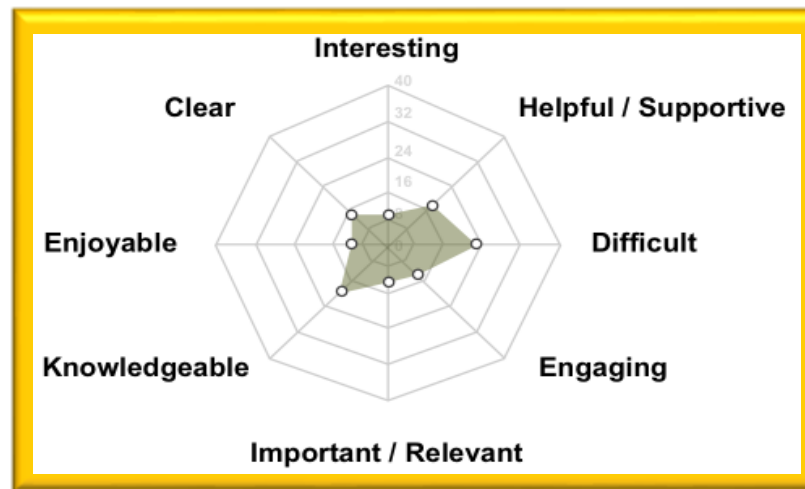
Very Good



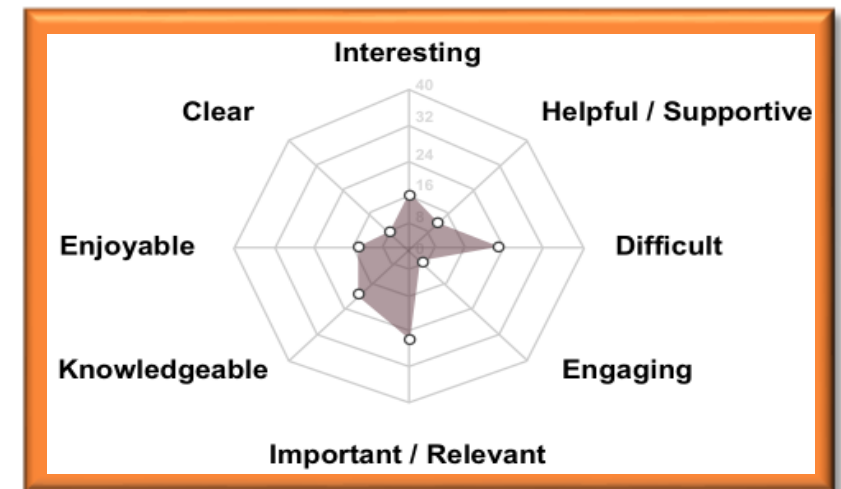
Good



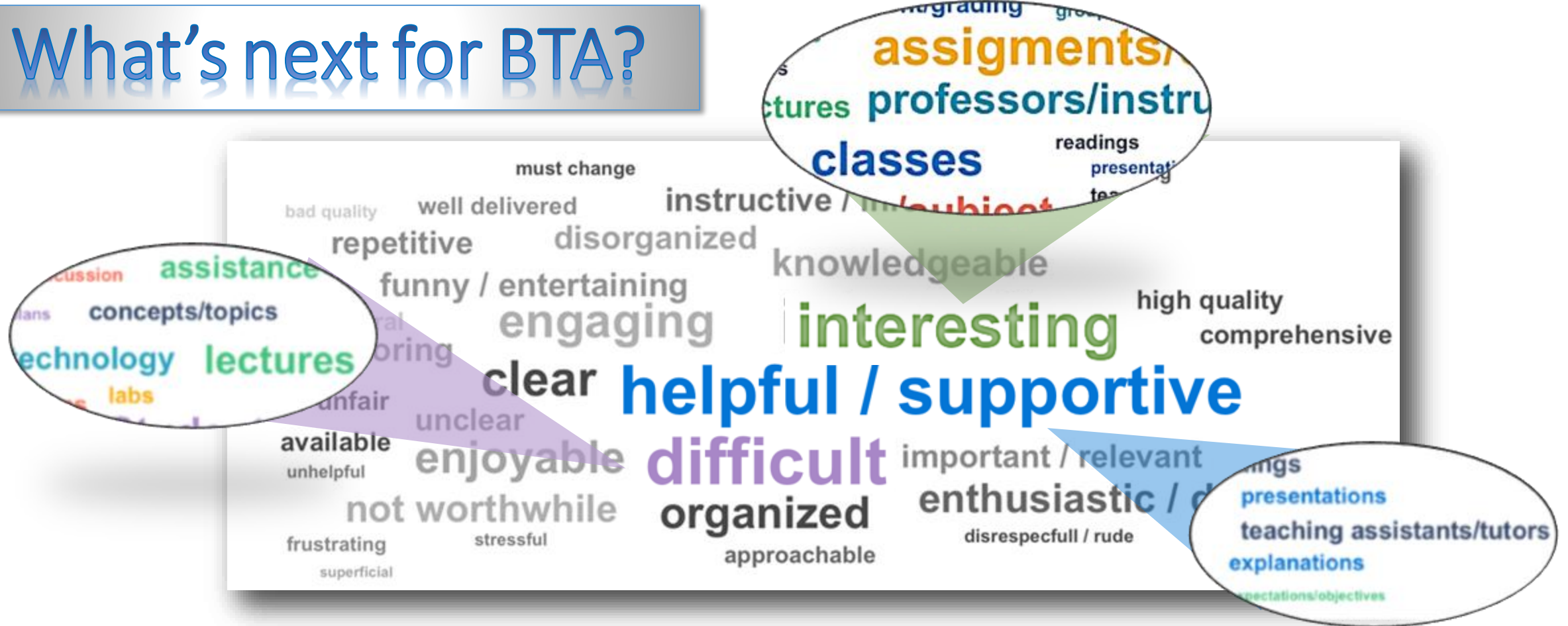
Fair



Poor



What's next for BTA?



What is *difficult*?

Who or what is *helpful / supportive*?

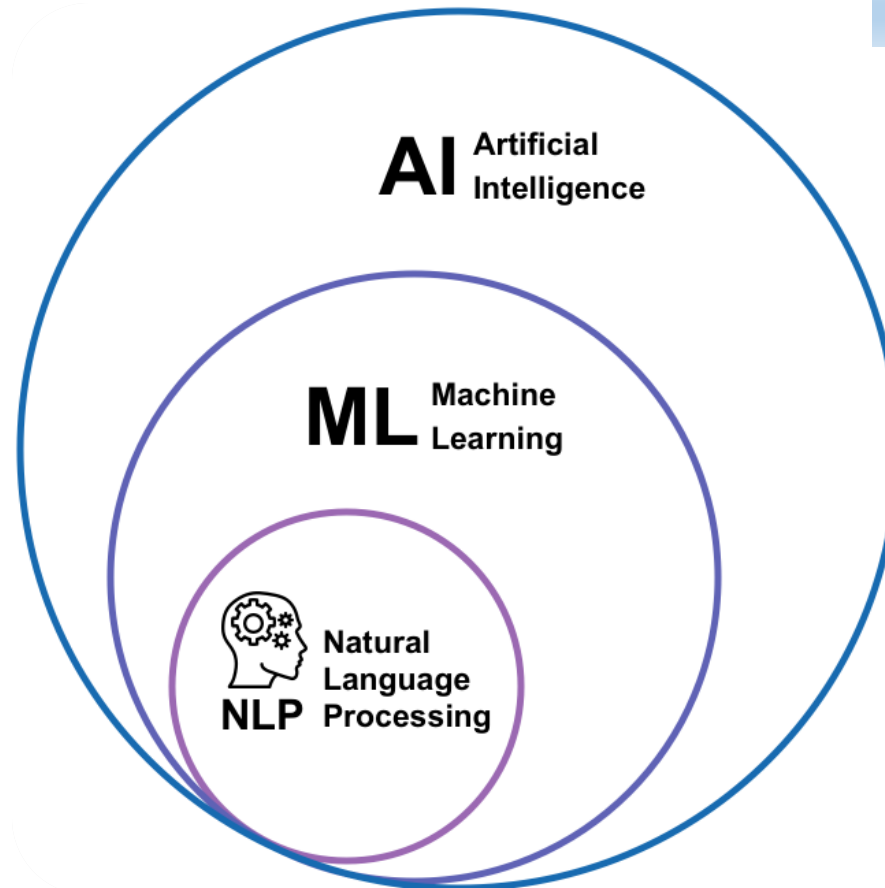
Natural Language Processing

How AI, ML and NLP work together

Machine Learning has the potential to detect patterns beyond hand written rules and validation.

We aim to improve upon handling of

- Misspellings
- Double and triple negations
- Indirect and passive phrases
- Regional differences in language



AI
Teaches system to do **intelligent** things

ML
Teaches systems to do **intelligent** things that can **learn** from experience

NLP
Teaches systems to do **intelligent things**, **learn** from experience and **understand** human language

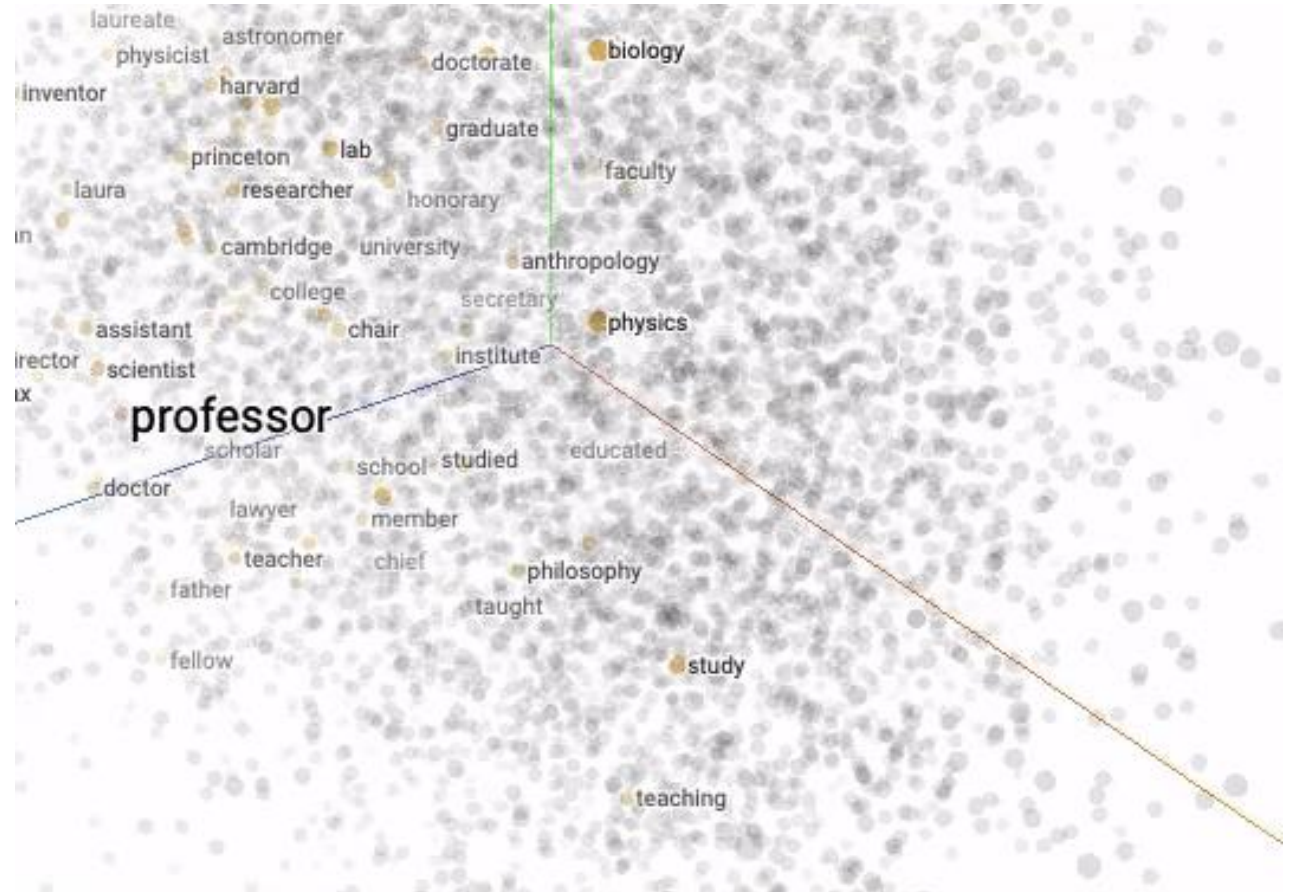
unintelligible text



Phase 1

Build the Language Model

- Goal of the language model:
To predict the next word in the sentence
- Approach: Unsupervised
- Computers don't understand words, they understand vectors. Similar words used in similar context will have similar vectors.

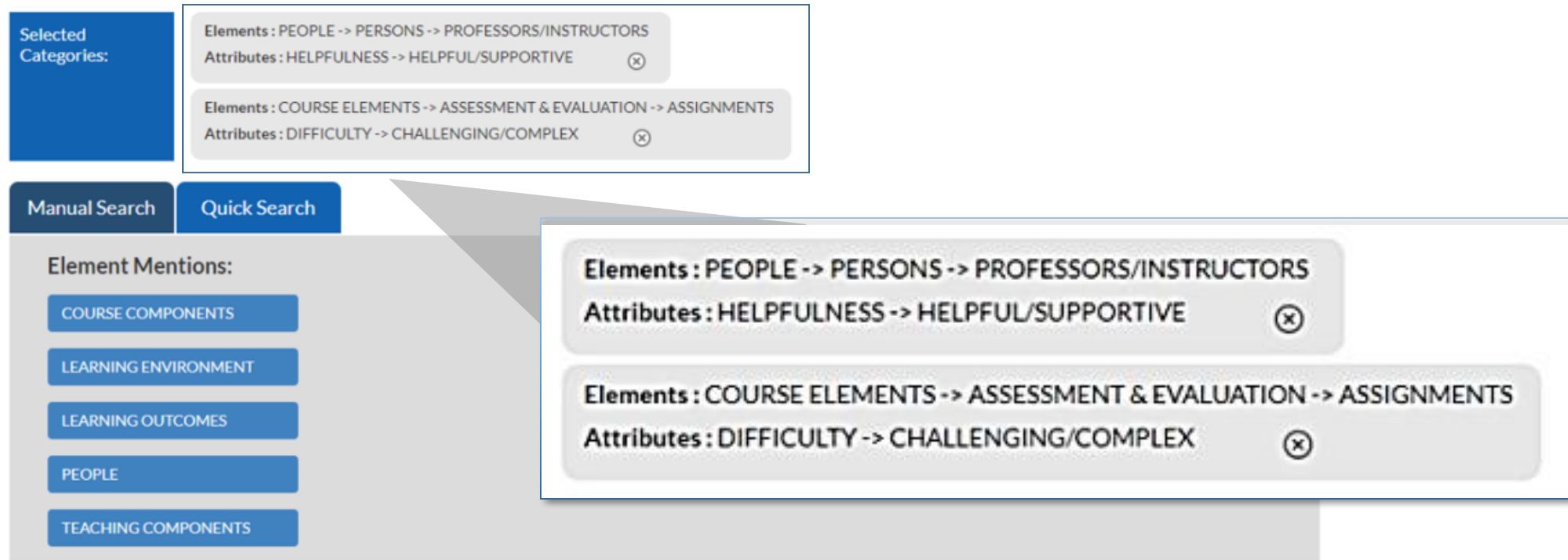


Detect The Themes

- Goal: We want the model to detect the sentiment and the themes found in comments related to Teaching & Learning experiences.
- Approach: Supervised
- Data collection:
 - Coming from the relevant fields, i.e. T&L
 - Containing regional and academic variations
- Data processing:
 - Manual tagging of the data with sentiment and themes
 - Divided into training and validation data sets
- Building of the model and the API

Tagging

Professor ~PERSON~ is always willing to help and even stayed after class a few times. He is super nice and clearly understands the material and presents in a way that helps me understand it as well. I was struggling with my assignment and he helped me work thought it.



The screenshot displays a tagging interface with the following components:

- Selected Categories:** A blue box on the left containing two category tags:
 - Elements : PEOPLE -> PERSONS -> PROFESSORS/INSTRUCTORS
 - Attributes : HELPFULNESS -> HELPFUL/SUPPORTIVE (with a close icon)
- Manual Search / Quick Search:** Two tabs at the top of the main panel.
- Element Mentions:** A list of five blue buttons: COURSE COMPONENTS, LEARNING ENVIRONMENT, LEARNING OUTCOMES, PEOPLE, and TEACHING COMPONENTS.
- Tagging Panel:** A large grey area on the right containing two tags:
 - Elements : PEOPLE -> PERSONS -> PROFESSORS/INSTRUCTORS
 - Attributes : HELPFULNESS -> HELPFUL/SUPPORTIVE (with a close icon)
 - Elements : COURSE ELEMENTS -> ASSESSMENT & EVALUATION -> ASSIGNMENTS
 - Attributes : DIFFICULTY -> CHALLENGING/COMPLEX (with a close icon)

1st Year Students

2nd Year Students

All Students



Now let's uncover:

- Top poor experiences
- Top good experiences

Let's uncover: What students mentioned by year of study

Analysis on: Qualitative Question
 + 'Overall what did you think of this course?'

1st Year Students



Now let's uncover:
 Now let's uncover:

- Top poor experiences
- Top good experiences



Program of Mechanical Engineering Results

What's next for BTA

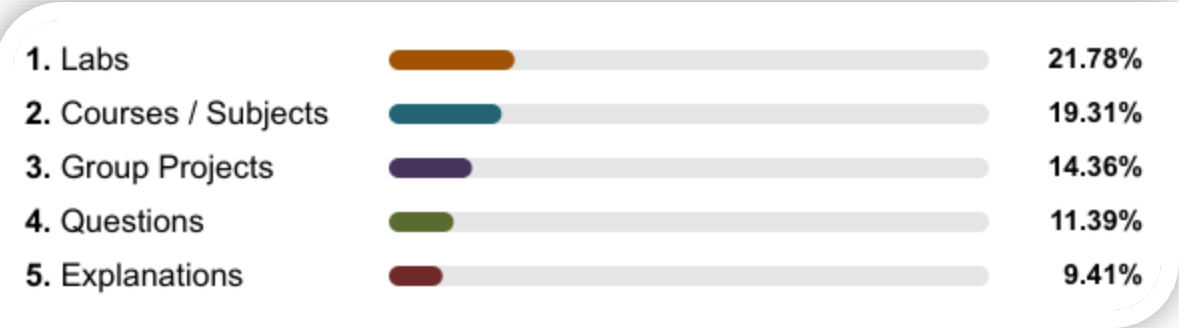
Analysis on: Qualitative Question

+ 'Overall what did you think of this course?'

1st Year Students



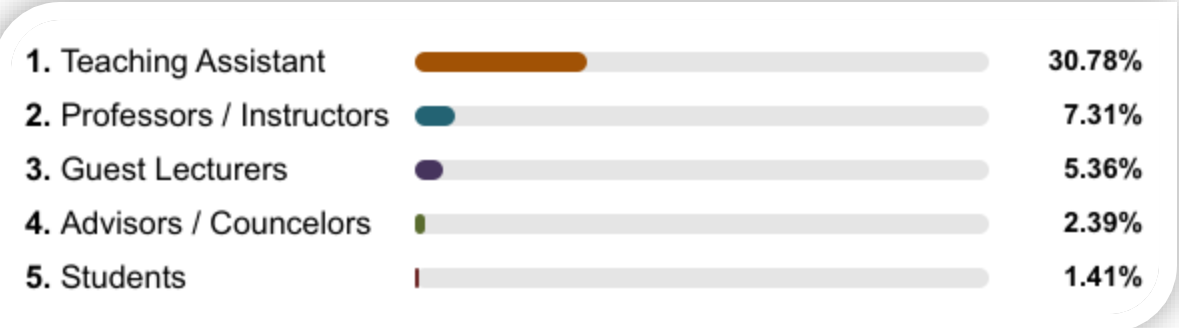
What was Difficult



Now let's uncover:

- Top 2 poor experiences
- What elements mentioned made these experiences poor

Who or what was Unavailable



Discovering themes automatically

- Goal: We want the model to find potential new themes automatically, and a committee will decide whether to include the newly discover themes.
- Approach: Semi- Supervised
- Unsupervised Learning is more difficult to implement since the machine doesn't have common sense. The patterns it will find will not necessarily make sense to a human. Hence the big challenges behind an unsupervised learning algorithm.



We need your help to train the model!

What we need:

- Open text responses from different regions and across different academic domains
 - Comments should be anonymized
 - Comment titles should be related to Teaching & Learning
 - Any names found in comments will be masked as much as possible
 - New data is needed yearly to update the model
- Your help with tagging the comments
 - We aim to have a minimum of 1000 tagged comments per theme
 - You will only access your own data if you help us tag the comments
- How this data will be handled, stored and processed for the project?
 - We will not share this data outside of Explorance
 - How long we will keep the data?

Thank you!

- Questions?
- *You can contact directly with lhua@explorance.com in case you are interested to contribute*