Unveiling the Blue Text Analytics machine learning project

- *Presented by : Zelbery Bedard*

explorance.

AGENDA

1. Overview of BTA concepts

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

Shift in habits...

explorance.

ABC UNIVERSITY
Course-Instructor Survey

Open-ended feedback

Quality Digital Quantity Contextual

Course Component Evaluation 2018 for 401791 African Cultures and Societies

explorance.

BTA Dictionary
Teaching & Learning Themes

What are BTA Themes?

- A grouping of terms and concepts
- Focus: Unveiling insights and improving Teaching & Learning experience
- Categories of Themes: Attributes, Elements Mentioned

Attributes
(Adjectives)

approachable, poor delivery, disrespectful / rude, unresponsive, keep as is, responsive, unkind / unpleasant, well delivered, enthusiastic / dedicated, expensive, kind / personable, important / relevant, clear, must change, unavailable, fair, looking knowledge, bad quality, funny / entertaining, frustrating, easy, helpful / suportive, stressful, neutral, engaging, unclear, available, interesting, high quality, unfair, boring, organized, difficult, instructive / informative, unapproachable, unhelpful, superficial, repetitive, knowledgeable, comprehensive, disorganized, notengaging, not worthwhile, enjoyable

Elements Mentioned
(Nouns)

teaching methods, tutorials/workshops, attendance, assessment/grading, group projects, discussion, assistance, assignments/exams, lesson plans, questions, concepts/topics, technology, lectures, professors/instructors, interactions, students, classes, readings, notes/handouts, labs, office hours, in-class work, comprehension, course/subject, presentations, teaching assistants/tutors, examples, feedback, explanations, instructions, content/materials, expectations/objectives, textbooks, practical aspects, learning atmosphere, guest lecturers

explorance.

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:
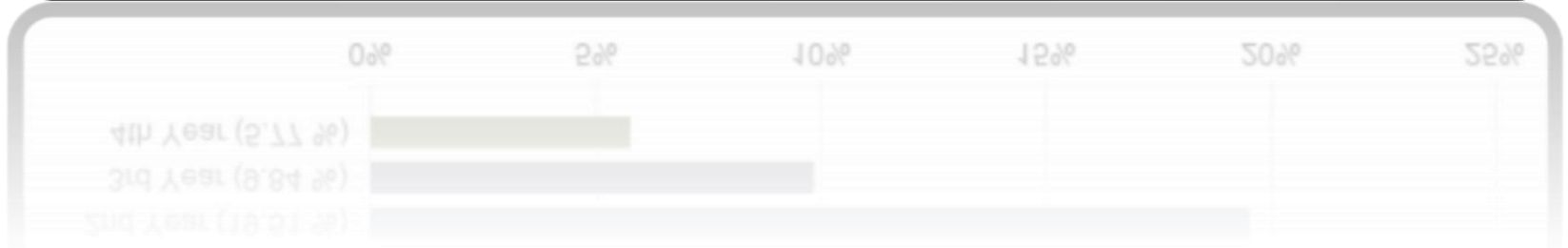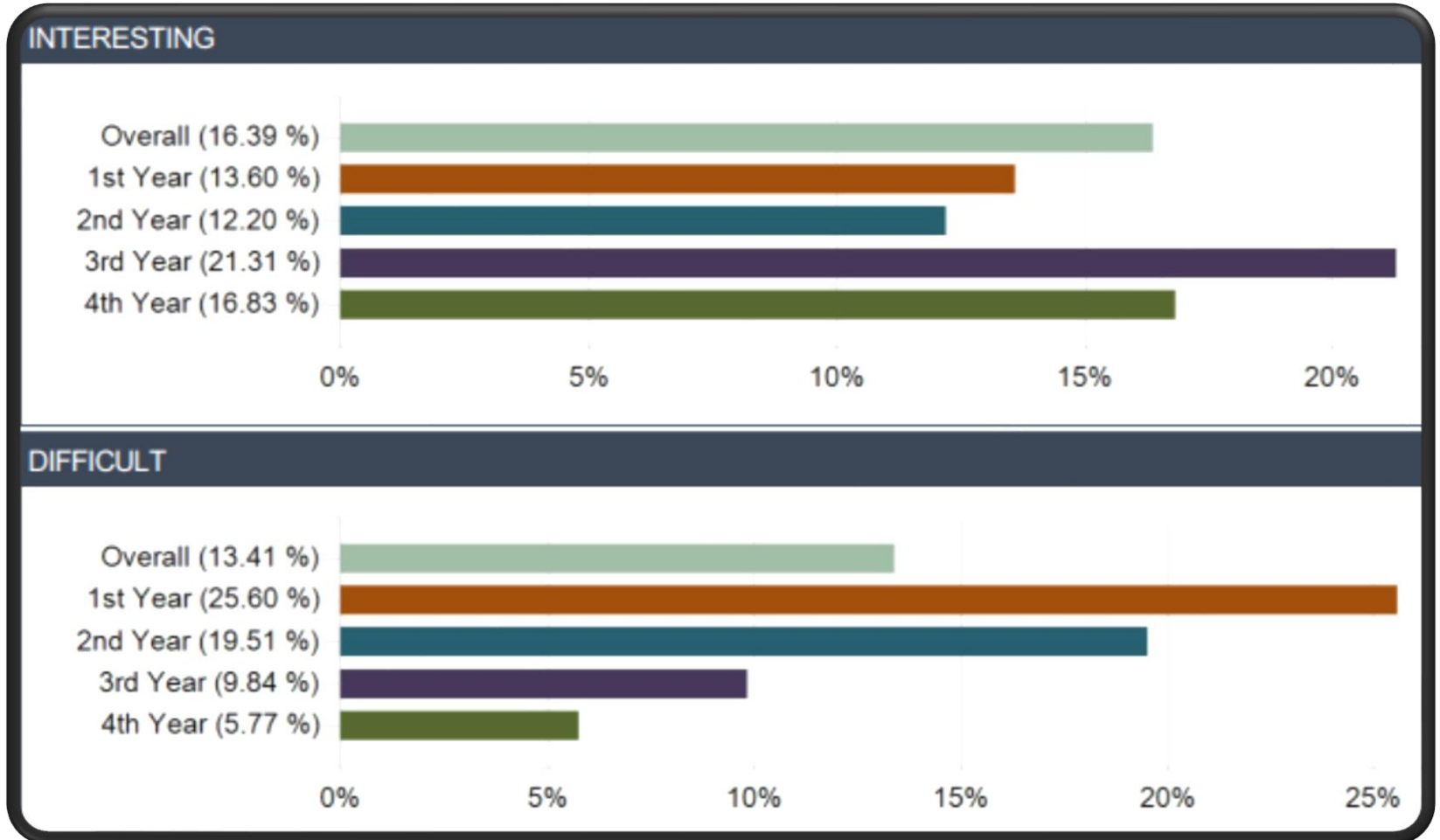accurate, clear, exact, explicit, precise, straightforward

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

explorance.

BTA Analysis
Themes compared across demographics

Comparing:

a) Top 2 *Attributes mentioned* in the program

b) Student year of study demographic data

INTERESTING

Overall (16.39 %)	
1st Year (13.60 %)	
2nd Year (12.20 %)	
3rd Year (21.31 %)	
4th Year (16.83 %)	

0% 5% 10% 15% 20%

DIFFICULT

Overall (13.41 %)	
1st Year (25.60 %)	
2nd Year (19.51 %)	
3rd Year (9.84 %)	
4th Year (5.77 %)	

0% 5% 10% 15% 20% 25%

explorance.

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

BTA Concepts
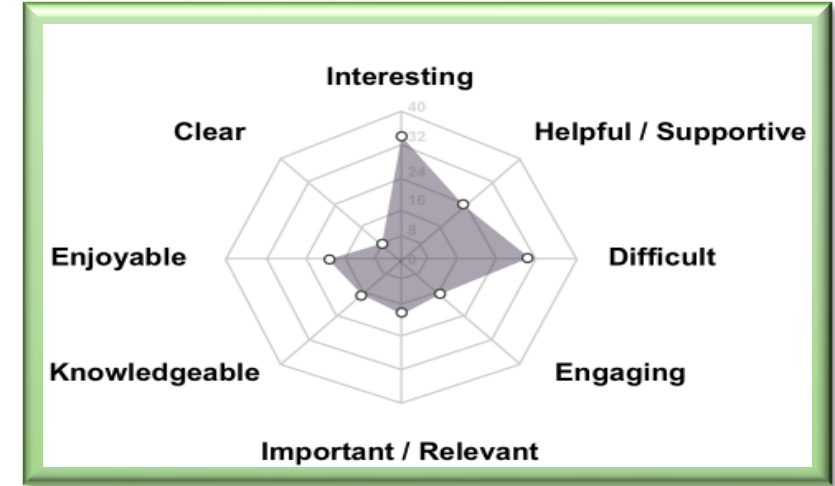
BTA Analysis
Themes compared across a rating question

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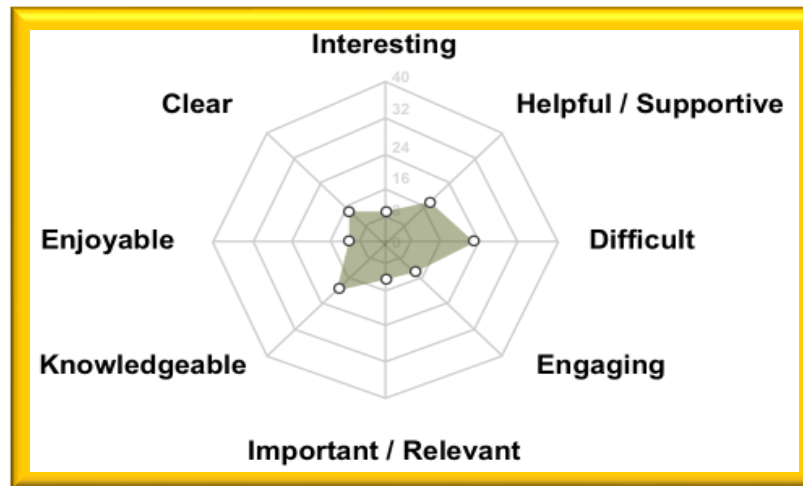b) "Overall what did you think of this course?" rating question
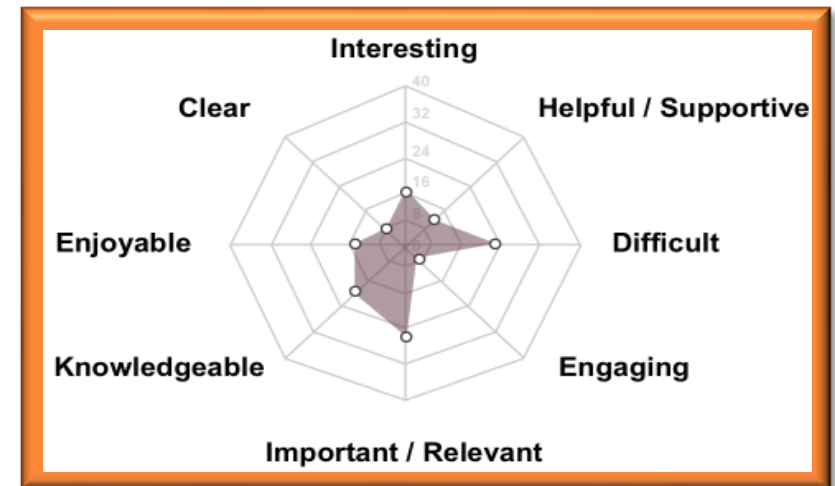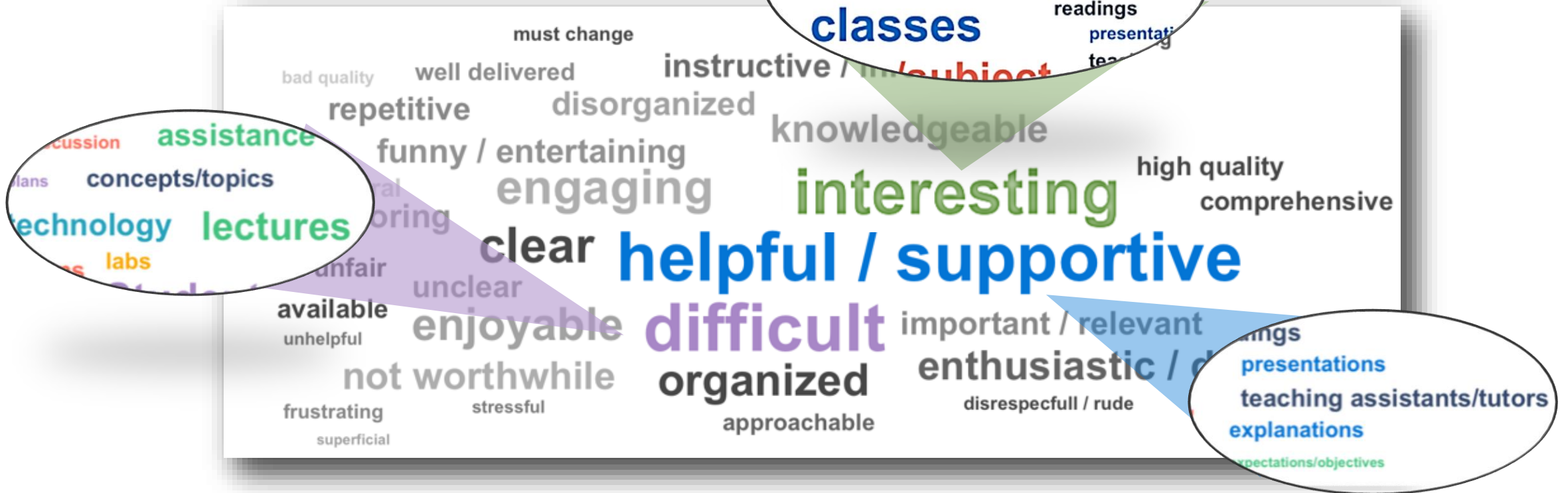
Very Good

Good

Fair

Poor

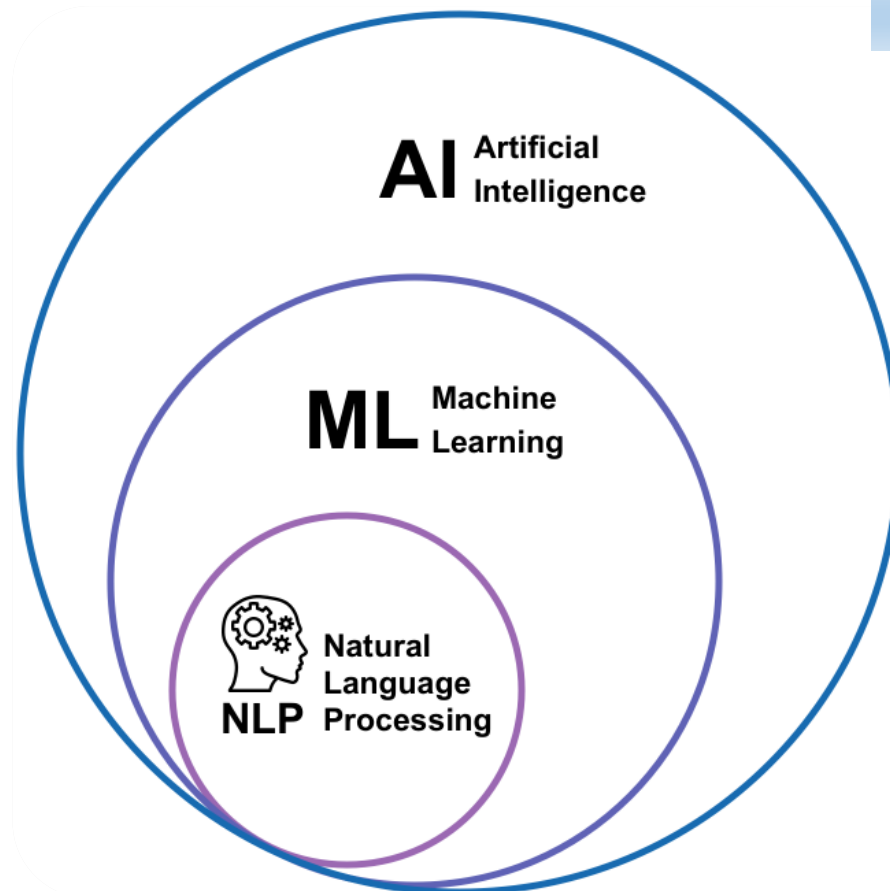# Natural Language Processing

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

How AI, ML and NLP work together

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

❖explorance.

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.

explorance.

Phase 2
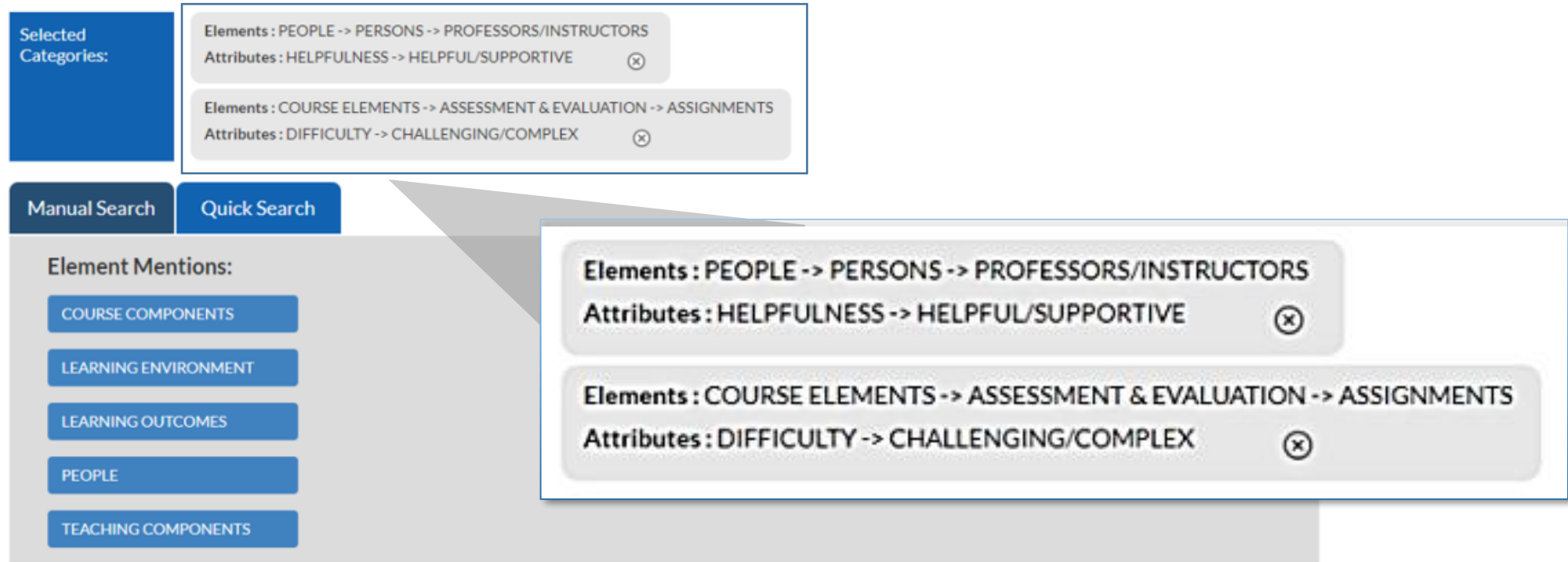
Detect The Themes

- 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 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.

1st Year Students

2nd Year Students

All Students

3rd Year Students

4th Year Students

Now let's uncover:
- Top poor experiences
- Top good experiences

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

explorance.

What's next for BTA

1st Year Students

bad quality
well delivered
repetitive
Enthusiastic
not engaging
must change
Organized
Not worthwhile
Enjoyable
high quality
unfair
neutral
Unresponsive
Important
easy
boring
comprehensive
To Much or too little
Difficult
Challenging
available
unclear
Unavailable
unhelpful
Helpful / Supportive
Right Amount
Approachable
Interesting
frustrating
stressful
superficial
disrespectfull / rude

Now let's uncover:
- Top 2 poor experiences
- What elements mentioned made these experiences poor

What was Difficult

1. Labs		21.78%
2. Courses / Subjects		19.31%
3. Group Projects		14.36%
4. Questions		11.39%
5. Explanations		9.41%

Who or what was Unavailable

1. Teaching Assistant		30.78%
2. Professors / Instructors		7.31%
3. Guest Lecturers		5.36%
4. Advisors / Councelors		2.39%
5. Students		1.41%

explorance.

Phase 3

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.

explorance.

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?

explorance.

Thank you!

- Questions?

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

www.explorance.com