Aalto University

Categorizing open response feedback with machine learning

Ville Kivimäki, Thomas Bergström, Jiri Lallimo, Alex Jung

Agenda

- Business need and background
- Machine learning solution and approach
- Conclusion

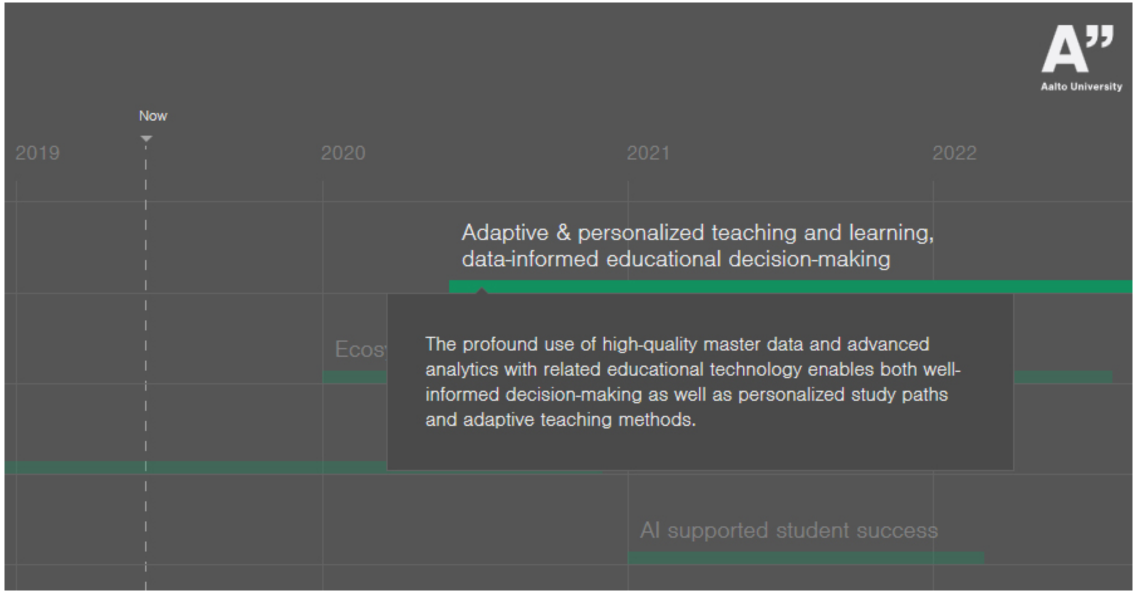
General interest in utilizing unstructured data

80 percent of business-relevant information originates in unstructured form

“The 80%/85%/90% unstructured figures come from, well, everywhere. “X percent unstructured” is so obvious to all of us working in text analytics that we’ve never dug up the research or even ascertained that it actually exists. This bit of common wisdom falls in the category, “if it didn’t exist, we’d have to invent it,” and perhaps we did. It is qualitatively true in the sense that it reflects our experience.”

- Seth Grimes, 2008

University strategy



Development actions

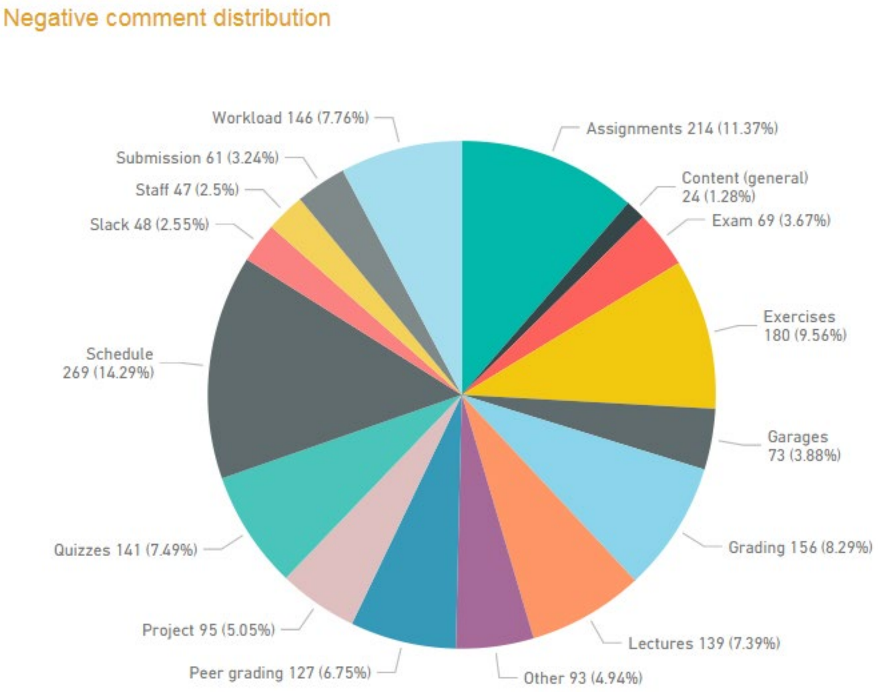
1. Attract and engage the best students through inspiring curricula that address societal challenges and offer multidisciplinary opportunities and consistently high quality learning experiences.
2. Develop and deploy forerunner digital learning solutions to improve learning outcomes.
3. Ensure success of our students' capabilities and wellbeing in a multicultural and diverse community by providing first-class study support.
4. Strengthen the working life capabilities of students by leveraging deep industry-academia relationships in educational projects, courses and internships.

Practical problem

”I have my hands tied in arranging a conference. At the same time I should be planning how to develop our course from last semester. I have thousands of feedback items collected from the last course. Could you help me to analyze the feedback?”

Background


course
Project
learning
machine
Lectures
Good
Interesting
assignments
Very
students
really
exercises
Data
lecture
more
methods
Well
work
nice
Slack
topic
Theory
theoretical
follow
one
Exercise
Lot
analysis
Practical
Topics
Great
helpful
Subject
Enough
Deadlines
Coding
applications
mycourses
implementation
introduction
mathematical
sessions
grading
motivation
Extremely
materials
algorithms
understand
liked
focus
lecturers
feedbac
experience
Gave
working
examples
quizzes
slides
Kaggle
Competition
Useful
helped
Easy
Motivated
lecturer
problems
peergrading
professor
additional
possible
people
everything
using
learned
assistants
teaching
knowledge
concepts
schedule
student
guest
approach
concrete
important
practice
without
tools
aspect
relevant
first
Those
questions
informative
different
Overall
Alex
through
doing
Garages
Better
help
ask
book
content


Negative comment distribution
Workload 146 (7.76%)
Assignments 214 (11.37%)
Submission 61 (3.24%)
Content (general) 24 (1.28%)
Staff 47 (2.5%)
Exam 69 (3.67%)
Slack 48 (2.55%)
Exercises 180 (9.56%)
Schedule 269 (14.29%)
Garages 73 (3.88%)
Quizzes 141 (7.49%)
Grading 156 (8.29%)
Project 95 (5.05%)
Lectures 139 (7.39%)
Peer grading 127 (6.75%)
Other 93 (4.94%)

Business opportunity

Course development
(teachers)

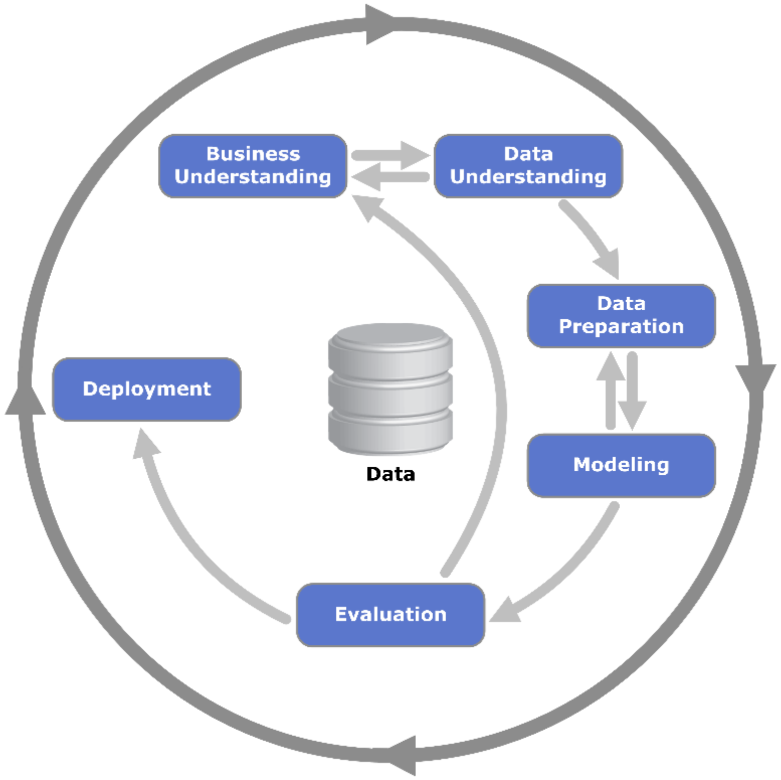
Programme development
(leaders)

University's quality
system
(funding)

Aalto University

Can machine learning help to categorize students' textual feedback?

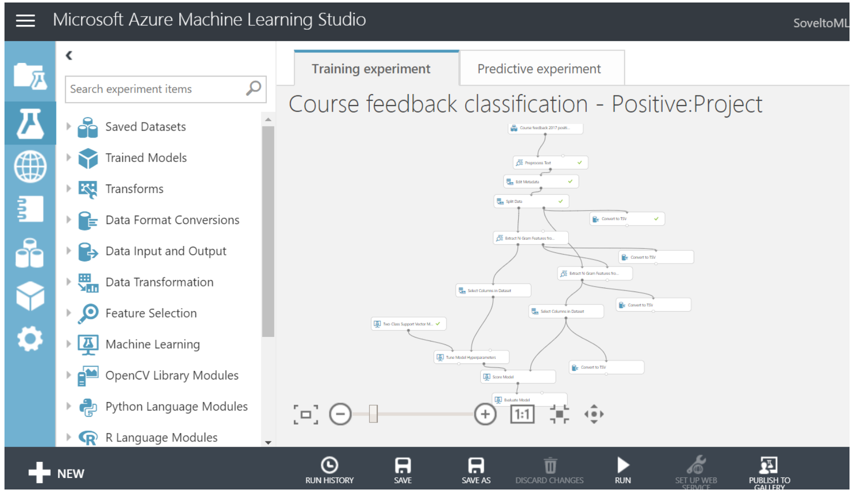
From business need to solution: How to create a text classification model for categorizing student course feedback?



Cross-industry standard process for data mining (CRISP-DM) is a common approach for analytics model creation - first 3 steps typically takes 80-90% of total time

Azure Machine Learning Studio was used to preprocess text and train the model

Azure Machine Learning Studio



- Data pre-processing (removing unnecessary characters and stop words, stemming)
- Feature extraction
- Model building and training
- Model evaluation
- Model deployment

Step 1: Data labeling

Student curse feedback

Labels (output)

	Positive feedback	Negative feedback	Positive: Assignments	Positive: Exercises	Positive: Lectures	Positive: Project	Positive: Garages	Positive: Schedule	Positive: Grading	Positive: Workload
2	I really liked the data-analysis project. It was	I feel that the HAs could have more focus	0	0	0	1	0	0	0	0
3	The lecture slides were great.	The project would've been even more us	0	0	1	0	0	0	0	0
4	Lots of useful extra material included in the le	The mycourses platform for the Quizzes	0	0	0	0	0	0	0	0
5	The lectures very clearly explained their topic	The schedule of the course was extremel	0	0	1	0	0	0	0	0
6	The lecturer was enthusiastic and a good spe	The amount of quizzes and the amount tl	0	0	1	0	0	0	0	0
7	The content was varied and extensive.	The assignments were often quite unclea	0	0	0	0	0	0	0	0
8	Extremely interesting course content.	Peergrading feels like a waste of time, I g	0	0	0	0	0	0	0	0
9	Skilled lecturers all around.	Emphasis of points from different tasks c	0	0	1	0	0	0	0	0
10	Project work assignment had a great topic.	Visiting lecturers should be at the end of	0	0	0	1	0	0	0	0
11	The course had a lot of visiting lecturers and	The workload divided a bit too unevenly.	0	0	1	0	0	0	0	0
12	It was nice to have project that actually is a b	The points distribution was quite uneven	0	0	0	1	0	0	0	0
13	The idea of exercise garages was great! How	Sometimes the HAs felt too much of a ma	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
14	friendly TAs	could've introduced and compared vario	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
15	decent material (slides)	at a point, I felt the course-work got too l	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
16	awesome lecturers	some assignment questions did not supp	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
17	The lectures were mostly informative, clear a	The home assignments were too flexible	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

Feedback data was extracted from Moodle LMS, 780 positive and 780 negative feedback

"Supervised learning requires labeled data, we had to manually label the data from Moodle LMS consisting of 780 positive and 780 negative feedback

Data understanding

Category	Count
Positive: Content (general)	206
Positive: Lectures	180
Positive: Project	140
Positive: Staff	67
Positive: Assignments	66
Positive: Slack	48
Positive: Schedule	34
Positive: Garages	25
Positive: Grading	25
Positive: Peer grading	20
Positive: Quizzes	9
Positive: Other	9
Positive: Workload	5
Positive: Exam	1

	Id	Ngram	DF	IDF
120	121	slide	26	1.326966
121	122	lecturer	28	1.294781
122	123	slack	28	1.294781
123	124	analysis	33	1.223425
124	125	material	33	1.223425
125	126	like	36	1.185637
126	127	data	37	1.173737
127	128	understand	40	1.139879
128	129	work	43	1.108471
129	130	help	44	1.098486
130	131	student	44	1.098486
131	132	machine	50	1.042969
132	133	exercise	52	1.025936
133	134	assignment	55	1.001576
134	135	topic	59	0.971087

Preparing the training data

	Original text	Pre-processed text
0	There is no exam. I learn best when I can actu...	exam ||| learn work | read ||| lecture slide
1	Various ways of collecting points for final grade	way point grade
2	Having doubt lectures and exercise garage was ...	doubt lecture exercise garage ||| attention in...
3	The lectures. I did not participate much but I...	lecture ||| participate like enthusiasm alex jung
4	Really interesting class where I learnt the ba...	class ||| learn basic ||| feel understand libr...
5	Having a lot of exercise sessions gave us flex...	lot exercise session flexibility
6	The project taught me a lot about using machin...	project teach lot machine library ||| tackle l...
7	Math. This course can be done wrong in two way...	math ||| course way ||| math | mean stuff lear...
8	Lectures were in generally good and informativ...	lecture | presentation addition ||| think idea...
9	great skills from the lecturer to keep attenti...	skill lecturer attention constant
10	General impression and knowledge around the co...	impression knowledge course topic | machine learn
11	Visitors in the lectures were inspiring.	visitor lecture inspire
12	The course delivered what it promised: an intr...	course deliver promise | introduction principl...
13	Learning about the concepts of Machine Learnin...	learn concept learn learn solve problem practice
14	A flexible schedule of assignments and the pro...	schedule assignment project
15	Good sequence between the readings with a natu...	sequence reading schema follow

Textual feedback has large number of noise such as stop words, abbreviations, typos, special characters – data pre-processing is key to high quality training data

Extracting information from text

Features extracted from the feedback

ocessed Positive ct.[read]	...	Preprocessed Positive aspect.[help]	Preprocessed Positive aspect. [student]	Preprocessed Positive aspect. [machine]	Preprocessed Positive aspect. [exercise]	Preprocessed Positive aspect. [assignment]	Preprocessed Positive aspect.[topic]	Preprocessed Positive aspect. [project]	Preprocessed Positive aspect.[learn]	Preprocessed Positive aspect. [lecture]	Preprocessed Positive aspect. [course]
.408248	...	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.408248	0.408248	0.000000
.000000	...	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
.000000	...	0.000000	0.000000	0.000000	0.577350	0.000000	0.000000	0.000000	0.000000	0.577350	0.000000
.000000	...	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.577350	0.000000
.000000	...	0.316228	0.000000	0.316228	0.000000	0.000000	0.000000	0.000000	0.316228	0.000000	0.000000
.000000	...	0.000000	0.000000	0.000000	0.707107	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
.000000	...	0.000000	0.000000	0.707107	0.000000	0.000000	0.000000	0.353553	0.000000	0.000000	0.000000
.000000	...	0.000000	0.316228	0.000000	0.000000	0.000000	0.000000	0.000000	0.158114	0.000000	0.316228
.000000	...	0.288675	0.000000	0.000000	0.288675	0.000000	0.000000	0.000000	0.000000	0.288675	0.000000
.000000	...	0.000000	0.000000	0.000000	0.000000	0.000000	0.000				
.000000	...	0.000000	0.000000	0.447214	0.000000	0.000000	0.447				
.000000	...	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000				
.000000	...	0.000000	0.000000	0.447214	0.000000	0.000000	0.000000				
.000000	...	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000				

Feature values for single feedback document

Features extracted from text using statistical methods (TF and TF-IDF) which shows how important a word is in a document

Select and train your model

Input data

ocessed Positive ct.[read]	...	Preprocessed Positive aspect.[help]	Preprocessed Positive aspect. [student]	Preprocessed Positive aspect. [machine]	Preprocessed Positive aspect. [exercise]	Preprocessed Positive aspect. [assignment]	Preprocessed Positive aspect.[topic]	Preprocessed Positive aspect. [project]	Preprocessed Positive aspect.[learn]	Preprocessed Positive aspect. [lecture]	Preprocessed Positive aspect. [course]
).408248	...	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.408248	0.408248	0.000000
).000000	...	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
).000000	...	0.000000	0.000000	0.000000	0.577350	0.000000	0.000000	0.000000	0.000000	0.577350	0.000000
).000000	...	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.577350	0.000000
).000000	...	0.316228	0.000000	0.316228	0.000000	0.000000	0.000000	0.000000	0.316228	0.000000	0.000000
).000000	...	0.000000	0.000000	0.000000	0.707107	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
).000000	...	0.000000	0.000000	0.707107	0.000000	0.000000	0.000000	0.353553	0.000000	0.000000	0.000000
).000000	...	0.000000	0.316228	0.000000	0.000000	0.000000	0.000000	0.000000	0.158114	0.000000	0.316228
).000000	...	0.288675	0.000000	0.000000	0.288675	0.000000	0.000000	0.000000	0.000000	0.288675	0.000000
).000000	...	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
).000000	...	0.000000	0.000000	0.447214	0.000000	0.000000	0.447214	0.000000	0.447214	0.000000	0.447214
).000000	...	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.707107	0.000000
).000000	...	0.000000	0.000000	0.447214	0.000000	0.000000	0.000000	0.000000	0.447214	0.000000	0.447214
).000000	...	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.832050	0.000000	0.000000

Model

Results

- Category A
- Category A
- Category B
- Category C
- ….

Textual data results in large number of features and a sparse data set - model needs to fit these characteristics

Evaluate model performance

Feedback category	Results
Positive: Lectures	Accuracy: 0.93, Precision: 0.92, Recall: 0.78
Positive: Assignment	Accuracy: 0.92, Precision: 0.81, Recall: 0.73
Positive: Project	Accuracy: 0.95, Precision: 0.91, Recall: 0.85
Positive: Content	Accuracy: 0.84, Precision: 0.75, Recall: 0.69
Positive: Staff	Accuracy: 0.95, Precision: 0.7, Recall: 0.7
Negative: Assignment	Accuracy: 0.83, Precision: 0.72, Recall: 0.6
Negative: Peer grading	Accuracy: 0.95, Precision: 0.82, Recall: 0.64
Negative: Lectures	Accuracy: 0.97, Precision: 0.79, Recall: 0.92

” Results indicate that machine learning can be used to categorize textual feedback - more data will improve model performance

Operationalize model

Deploy

Integrate

Automate

Scale

Monitor

Improve

Conclusions & lessons learned

- Important to utilize also the information in textual feedback, augments the structured feedback data
- University wide solution requires that we standardize the course feedback metrics, categories and processes
- Collect more training data from courses to improve model performance
- Think about how to operationalize the model in order to achieve business benefits – e.g., integrate with feedback and LMS systems
- Proof-of-concept solution has helped Aalto University to build machine learning competence in machine learning and text classification and get realistic understanding about the opportunities

Thank you!

Ville.Kivimaki@aalto.fi
Thomas.Bergstrom@aalto.fi