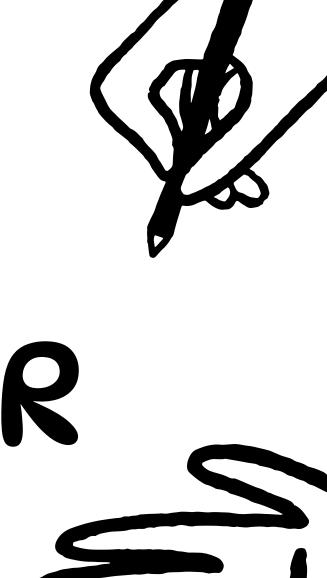


AUC COURSE RECOMMENDER

Adrianna Zaranska, Barbara Plebanowicz, Leon Wloch





Introduction

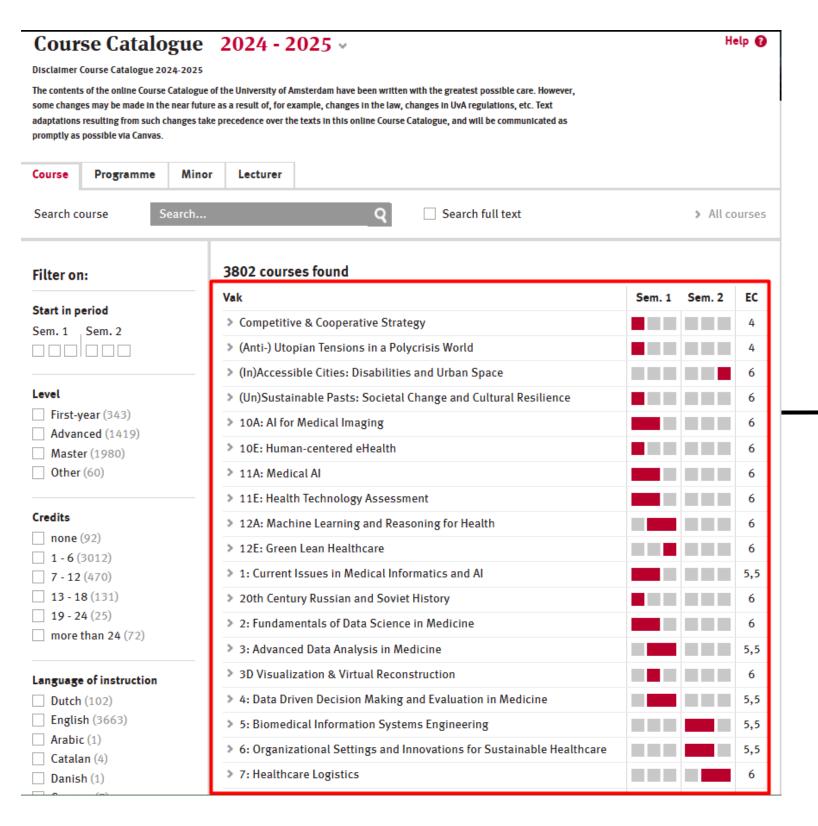
- Own experiences in AUC
- Struggle choosing courses in Liberal Arts and Sciences Programme
- Available dataset

	Health	Cognition	Economics	Environmental Economics and Policy	Law	International Relations	Political Science	Anthropology	Sociology	
			The Art Market and Culture Industry **		Artificial Intelligence, Technology and the Law **	International (dis)order: Past and Present **		Queering Media Studies **		
			Market Failures *		Corporate Social Responsibility *	Peace Lab (Rwanda * / Kosovo **)		Cultural Memory Studies *		
300			Advanced Macro-Economics **	Case Studies in Energy, Climate and Sustainability **	Global Environmental Governance	Diplomacy Lab *, **		Photograph as a Socio- Political Document *	Violence and Conflict	
30	Lifestyle and Disease **	Social Cognition **	Advanced Micro-Economics *	Urban Environment Lab **	Criminal Justice Systems **	Global Environmental Governance	International (dis)order: Past and Present **	Decolonisation in Historical Perspective *	Migration, Integration and Diversity *	
	Addiction **	Mind Reading: Multivariate Pattern Analysis **	Development Through an Equity Lens **	Global Environmental Governance	European Union Law *	International Crimes	Comparative Public Policy **	Religion, Trauma and Violence *	Contemporary Sociological Thought *	
	Human Stress Research *	The Empathic Brain *, **	Behavioural Economics *, **	International Sustainable Development **	Legal and Social Philosophy *	Violence and Conflict	Political Communication and Data Analytics *	Urban Anthropology Lab **	Race, Class and Gender Intersectionality *	
				Advanced	d Research Methods and Stat	tistics **				
					Ethics			Digital Habits, Digitized Lives *	History Lab *	
			Fundamentals of Macro- Economics *		International Law **		The History of Ideas **	Journalism *	Philosophy of Science *	
	Brain and Cognition **	Brain and Cognition **	Fundamentals of Micro- Economics *	Risk Management and Natural Hazards *	Human Rights Law and Politics *	Ethics	Empire and Its Afterlives **	World Religions *	Gender and Sexuality **	
200	Nutrition and Health **	Developmental Psychology **	Econometrics **	Environmental Law and Policy *	Constitutional and Administrative Law **	Human Rights and Human Security	Data Futures Lab **	Poetics of Protest **	Inequality and Poverty **	
	Epidemiology *	Cognition Lab **	International Political Economy **	Introduction to Geographic Information Systems *	Environmental Law and Policy *	Empire and Its Afterlives **	Democracy in Crisis *	Medical Anthropology **	Nations, Nationalism and Modernity **	
	Gastronomy: the Applied Sciences of Cooking *	Cognitive Psychology	Game Theory	Sustainable City **	Principles of Private Law *	Human Rights Law and Politics *	The Politics of Modernity *	Anthropologies of Community **	Sociology of the Other *	
		Statistical Methods for S	Social Sciences Research							
	Introduction to Health and Wellbeing *	Linguistics **	Perspectives on Economic Thought **	Introduction to Climate and Sustainability *						
100	Introduction to Public Health	Artificial Cognition: Pattern Recognition	Challenges of Food and Nutrition Security *	Introduction to Environmental Sciences						
10	Health, Resilience, and Human Flourishing *	Psychology	Economic Thought in a Historical Perspective *	Environmental Economics **	Law, Society and Justice	International Relations Theory and Practice	Classical and Modern Political Thought	Classical and Modern Anthropological Thought	Classical and Modern Sociological Thought	
	Mathematical Methods for Economics Methods for Social Science Research									

		Literature	Film	Philosophy	History	Culture	М	ledia	Art History	
		Existentialism in Literature and Philosophy **		Existentialism in Literature and Philosophy **			Visual (Culture **		
				Mathematical Logic *		More than Human *				
	300			Legal and Social Philosophy *		Race, Class and Gender Intersectionality *				ω
ř				Debates and Dialogues in Philosophy **		Cultural Studies of Affect and Emotion *				00
				Ancient Philosophical Texts **		Urban Anthropology Lab **				
	Contemporary Postcolonia Literature *		Film Philosophy *		Cultural Men	nory Studies *	Queering Me	edia Studies **	The Art Market and Culture Industry **	
		Advanced Creative Writing **	Film and the Body **	Modern Philosophical Texts *	Decolonisation in Historical Perspective *	Religion, Trauma, and Violence *	Media / Er	nvironment *	Photograph as a Socio- Political Document *	
		Poetics of Protest **								
		Creative Writing **	Film Lab *	Philosophical Logic *						
,	Author in Context *		National Cinemas *	Philosophy of Science *	History Lab *	Cases in Cultural Analysis **	Media	a Lab **	Bodies on Display *	20
	70	Literary Ecologies * Documentary *		The History	of Ideas **	Gender and Sexuality **	Journ	nalism *	Making of a Painting **	00
		Modernism and Postmodernism *	Film Analysis *	World Religions *	Nations, Nationalism and Modernity **	Sociology of the Other *		bits, Digitized ves *	Contemporary Art **	
		Adaptation Studies **		Ethics * Counterculture **		Perspectives on Games **		Portraiture and the Body *		
	Methods in the Humanities 2									
	Music & Text *								Intro to Design, Architecture and Urbanism **	
	Literature Off the Page **		re Off the Page ** Introduction to Philosophy II **		Early to Modern History **				Introduction to Art History 2 **	
ş	ğ	Introduction to Literature			A Golden Age? History and Heritage of the Dutch Republic * Introduction to Cultural Analysis **		Introduction to Media Studies		Introduction to Art History 1 *	00
	Methods in the Humanities 1 *									
İ		Sociology	Mathematics Info	mation Physics	Chemistry E	arth and Environment	Biology	Biomedica	al Health	
			Quantum Information and	Quantum Communication **		Human Evolution **		Life	estyle and Disease **	

		Mathematics	Information	Physics	Chemistry	Earth and Environment	Biology	Biomedical	Health	
		Quantum Information and Quantum Communication **			Human Evolution **			Lifestyle and Disease **		
				Case studies	in Energy, Climate and Sus	tainability **	Challenges in Health and Society *		ty *	
	۰				Advanced Research Me	ethods and Statistics **				
nd	30	Text Mining **			Advanced Geosciences **		Infectious Diseases **		Mind Reading: Multivariate Pattern Analysis **	30
cal		Discrete Mathema	atics and Algebra *			Urban Environment Lab **	Epigenetic Regulations **	Clinical Neurosciences **	The Empathic Brain **	
er		Mathemati	cal Logic *	Nanoscience **	Atmospheric Sciences **		Cancer Biology and Treatment *	Cardiovascular Diseases *	Addiction **	
		Partial Differential Equations *	Modelling Real World Problems **	Mathematics of Physics **	Molecular Sustainability **	Climate Sciences: Past and Present *	Conservation and Restoration Biology *	Neuroscience *	Human Stress Research *	
ı		Numerical Mathematics **	Information Lab **	Physics Lab **	Pharmacology **	Field Course in Environmental Earth Sciences **	Urban Ecology Lab **	Molecular Techniques Lab **	Health Lab **	
	Comple		Complexity Lab **		Chemistry Lab *		Cell Biology and Physiology Lab **			
7		Probability and Statistics **	Advanced Programming **		Making of a Painting **			Genes, Bioinformatics and Disease **	Gastronomy: the Applied Sciences of Cooking *	1
-	200	Game Theory	Maker Lab **		Medicinal Chemistry **	Hydrology and Watershed Management **	Game Theory	Metabolic Biochemistry **	Nutrition and Health **	200
nd		Philosophy of Science *	Philosophical Logic *	Statistical Mechanics *	Environmental Chemistry/ Eco-Toxicology *	Introduction to Geographic Information Systems *	Freshwater and Marine Biology **	Hormones and Homeostasis **	Medical Anthropology **	
		Dynamical Systems *	Machine Learning *	Quantum Physics *	Organic Chemistry *	Risk Management and Natural Hazards *	Molecular Cell Biology *	Human Body - Anatomy and Physiology II *	Epidemiology *	
		Vector Calculus *	Data Structure and Algorithms *	Thermodynamics *		System Earth *	Evolution and Origin of Human Diseases *	Immunology *	Brain and Cognition **	
					Life, Earth a	nd Universe *		Health, Resilience and Human Flourishing *		
	100	Linear Algebra	Intermediate Programming: Principles and Practise *	Introduction to the Energy Transition *		Introduction to Environmental Sciences	Ecology - from Soil to Society **		Challenges of Food and Nutrition Security *	
		Statistics for Sciences	Programming Your World	Electricity and Introduction to Climate Magnetism **		e and Sustainability *	Introduction to Biology *	The Human Body – Anatomy and Physiology	Introduction to Public Health	100
		Calculus	Artificial Cognition: Pattern Recognition	Introduction to Introduction to Physics * Chemistry		Introduction to Geological Sciences **	Introduction to Health and Wellbeing *		eing *	
										$\overline{}$

Dataset



Selenium

BS4

Text Mining



Objectives

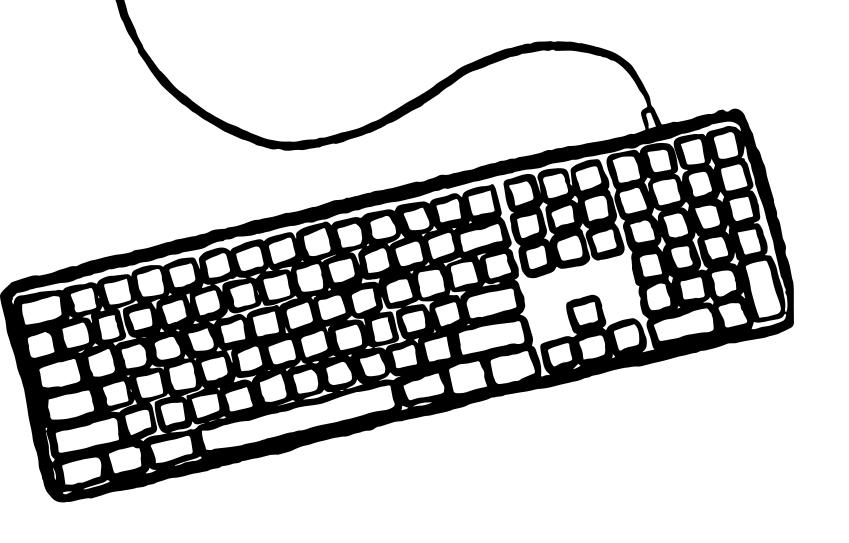
- 1 The student develops an understanding of the basic concepts and techniques used to mathematically represent and model language, including their underpinnings in linguistics.
- 2 The student develops programming skills for automatically extracting and processing textual data from different sources (such as textual corpora and social media).
- 3 The student is familiar with the main approaches to distributional semantics, including modern deep learning techniques, and is able to automatically extract distributional information from corpora.
- 4 The student is able to automatically characterise the semantic content of a document.
- **5** The student develops programming skills necessary for building his/her own text mining application (e.g., a search engine or a recommender system) by exploiting statistical learning approaches.

Contents

Constellation: Digital Worlds

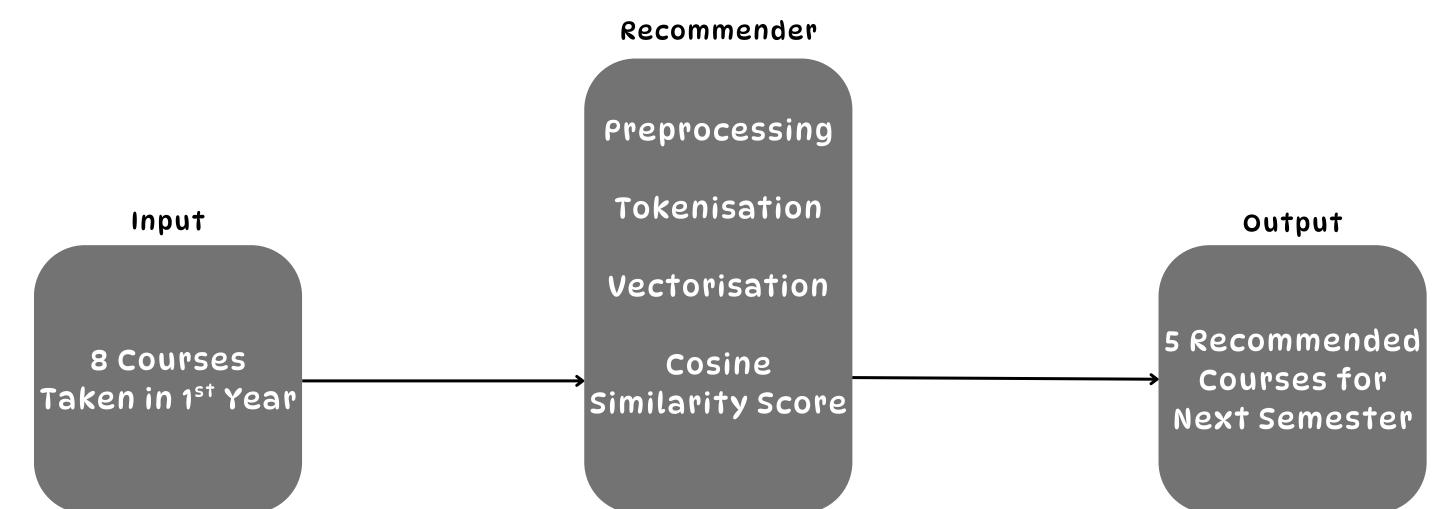
The increasing amount of digitally available text data contains a wealth of information about topics, people, products, behaviours. Due to its numerous applications (research, industry, government, non-profit, etc), extracting information from, and working with, texts is of crucial importance. This is the aim of text mining.

This course provides an introduction to text mining and natural language processing in Python, including real-world applications. The course will introduce fundamental concepts and techniques from



Recommender Overview

- Memory based recommender
 - Store and use vectors directly
- Content-based filtering
 - Regression



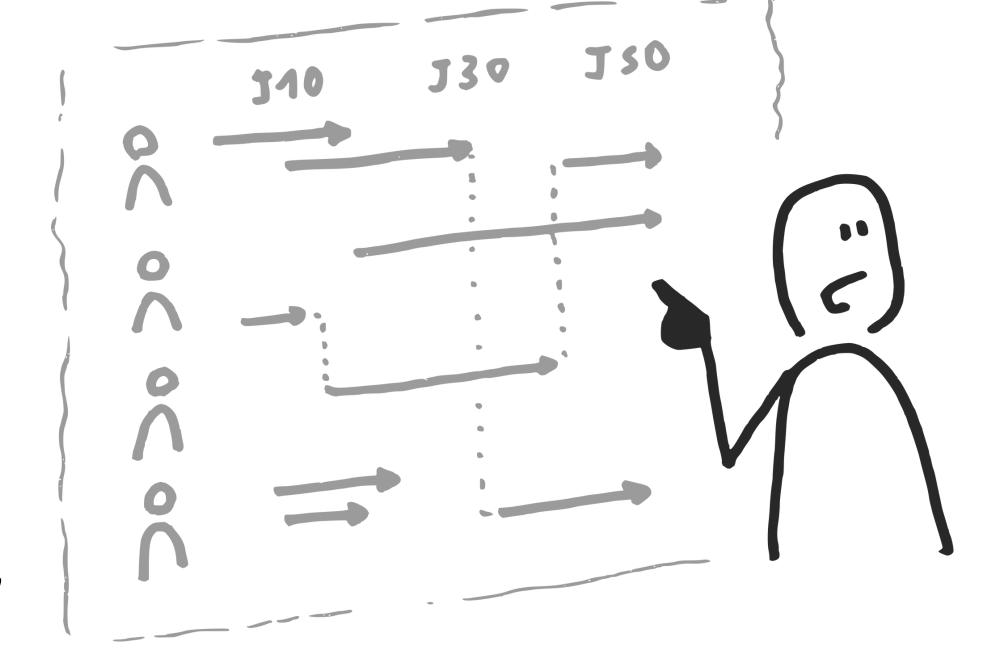


Methods

- Preprocessing:
 - o courses with no description
 - stop words
 - punctuation
- Tokenisation: NLTK tokeniser
- Vectorisation: Doc2Vec
- Similarity Measure: cosine similarity

Evaluation

- Accuracy Metrics*:
 - o Precision @ K
 - Normalized Discounted
 Cumulative Gain (NDCG)
- Diversity Metrics:
 - o Intra-List Diversity (ILD) @ K
 - Trade-off Between Diversity and Accuracy



*we would be grateful if you could share your courses from previous semesters;))

$$ILD@k = (\sum_{i=1}^{k+1} \sum_{j=1}^{k+1} d(i,j))/(k*(k-1))$$

Example Results

art history/history

	course_name	course_catalogue_number
1357	[history, lab]	9002601HUY
2537	[race, class, and, gender, intersectionality]	900374SSCY
2044	[migration, citizenship]	7332B005AY
2372	[postcolonial, encounters, in, arts, and, cult	111221596Y
1635	[journalism]	900258HUMY

physics

	course_name	course_catalogue_number
2209	[optimisation, of, business, processes]	53348OPB6Y
2274	[philosophy, of, science]	5354PHSC6Y
2142	[natural, resource, economics]	6414M0505Y
2604	[research, masters, internship, linguistics, and]	176527036Y
1703	[literary, ecologies]	900202HUMY

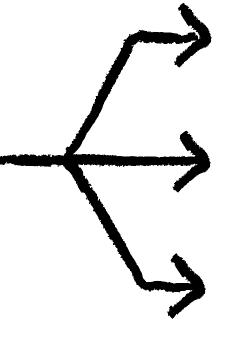


Limitations

- Short course descriptions
- No course descriptions for some courses
- Can't utilise collaborative-based filtering : no user data
- AUC dynamics: every year changes in course availability and programme structure

01. create a Graphical User Interface (GUI)

Our Extensions



O2. Incorporate time periods of the courses



Get real student data and evaluate the recommender.

Future Possible Extensions



- Extending the recommender to VU courses
- AUC graduation requirements: including requirements academic core, electives, other major requirements
- Extra input: interests, making the recommender for choosing courses from other majors



Thank you!



Questions?

