Project 6 Topic Modeling

Part 1 - Topic Modelling using LDA

1.1 For each time-period assign a name to each generated topic based on the topic's top words. List all topic names in your report.

Period 1: before 1990

Nr	Keywords	Topic Name
0	theory problems algorithms simulation decision parallel application solution applications optimal digital control	incoherent
1	computer logic model programs digital performance design networks using applications systems simulation	incoherent
2	problem programming optimal language digital processing software research solution parallel linear control	Parallel Programming
3	data method network models application languages solution processing problem using analysis programming	Computer Networks
4	note information linear functions applications finite technical programming systems time problem decision	Linear Programming
5	algorithm design analysis approach sets performance new using implementation parallel linear digital	Algorithm Design
6	systems using parallel performance implementation decision distributed linear control digital design processing	Distributed Systems
7	control networks new recognition distributed time pattern optimal systems approach digital linear	incoherent

Period 2: from 1990 to 2009

Nr	Keywords	Topic Name
0	time algorithm linear new network models algorithms efficient high robust equations management	Algorithms
1	method study problems evaluation case space programming equations finite linear new performance	incoherent
2	design approach nonlinear optimal fuzzy modeling computing robust control equations new time	incoherent
3	based control model methods computer robust time linear simulation network detection dynamic	Computer Simulation
4	using analysis networks performance problem multi dynamic neural wireless mobile recognition network	Mobile Computing
5	systems data information multiple digital linear robust management time control nonlinear analysis	Information Management
6	adaptive application structure non theory knowledge scheme management robust linear control finite	Knowledge Management
7	learning estimation applications order image distributed graphs web software power real development	incoherent

Period 3: from 2009

Nr	Keywords	Topic Name
0	systems networks wireless performance novel evaluation sensing sensor communication linear distributed power	Wireless Networks
1	based detection design linear mobile recognition prediction computing images cloud feature method	Image Recognition

Nr	Keywords	Topic Name
2	data method network optimization application applications modeling equations methods hybrid cloud problems	Cloud
3	adaptive nonlinear information models framework energy deep management social scheduling systems tracking	incoherent
4	control model estimation robust sensor selection tracking systems linear nonlinear stochastic distributed	Distributed sensors
5	learning neural efficient optimal distributed scheme problem multiple problems stochastic machine deep	Machine Learning
6	approach algorithm study new time algorithms online tracking case improved research optimization	Optimization
7	using analysis image dynamic power classification fuzzy communication machine feature learning selection	Classification

1.2 Do the topics make sense to you? Are they coherent? Do you observe trends? Discuss in 4-6 sentences

In our opinion the given topics only somewhat make sense since they are not very coherent. Therefore, between different topics there is always a certain overlap and various keywords appear in multiple topics at a time (e.g. digital, parallel, linear). This overlap between the sets of keywoards makes the topics more ambigous and blur the lines between topics. Therefore, it is often very hard to find a suitable name for a topic. Often a set of keywoards had to be labeled as **incoherent**.

Part 2 - Resource Limited Competition: Sentiment Analysis

2.1 Assign a name to each topic based on the topic's top words (for each time-period). List all topic names in your report.

Period 1: before 1990

Nr	Keywords	Topic Name
0	computer review introduction research software network new simulation graphics computers technology artificial book science report intelligence operations architecture communication local	Computer Network
1	note problem problems technical programming solution editor letter optimal scheduling networks solving times dynamic queues linear queue decision allocation integer	Programming
2	system data design information distributed processing language database management expert structures chemical retrieval development interactive base systems online knowledge structure	Database System
3	algorithm parallel algorithms sequential binary fast machines using circuits matrix detection method trees fault switching tree search efficient computing transform	Parallel Computing'
4	sets der graphs und von automata finite zur properties ein fuumlr boolean recursive uumlber die degrees automaten arithmetic eine relations	Graph Theory
5	control analysis systems recognition pattern adaptive using model linear optimal approach application estimation speech classification digital identification automatic image multivariable	Control System
6	timing vehicle augmented biological multiprogramming combined priorities texts references usage disk reconfigurable magnetic emergency urban balancing signature uncertain secondary interference	Multiprogramming
7	logic theory de languages theorem modal propositional semantics calculus set logics order symbolic grammars meeting association completeness first natural programs	Computational Linguistics

Period 2: from 1990 to 2009**

Nr	Keywords	Topic Name
0	problems problem equations solution finite solutions numerical equation methods method solving order differential approximation generalized boundary functions convergence difference value	Programming
1	theoretic rates spatially serial arrival modes simplified utilizing membership window composite cross underwater stage equivalent absolute replacement variations various transient	Memory Management
2	information development case web knowledge study technology management research electronic system support collaborative paper software health business framework process online	Internet technology
3	networks wireless performance sensor mobile routing scheduling protocol network traffic scheme access power distributed qos allocation atm dynamic communication packet	Computer Networks
4	systems control robust stability linear design adaptive uncertain nonlinear output feedback controller stabilization optimal discretetime state timevarying controllers approach delays	Control System
5	special issue introduction computer editorial de logic intelligence book language guest semantics section review amp isbn applications advances pp conference	Digitalization
6	analysis data model fuzzy neural models molecular approach classification prediction network gene modeling mining genetic application selection protein decision structure	Neural Networks
7	image images detection recognition segmentation using compression estimation brain speech imaging face magnetic radar reconstruction matching motion automatic sar signals	Computer Vision

Period 3: from 2009

Nr	Keywords	Topic Name
0	adjustment selforganizing window nested simplified weighting train progressive optimum cross partition overlapping fractal impulse employing pairwise multi redundancy polar transformer	Machine Learning
1	analysis surface data imaging land models temperature using water estimation series forest satellite mapping modeling soil radar comparison field cover	Big Data
2	optimization fuzzy algorithm multiobjective problem decision swarm approach genetic evolutionary model hybrid particle new problems selection search programming set algorithms	Optimization
3	image deep learning detection recognition classification neural feature segmentation images representation network face convolutional object features based sparse machine fusion	Deep Learning
4	networks wireless sensor power energy cognitive cellular allocation interference spectrum channels channel radio protocol relay massive mimo ad cooperative communication	Networks Science
5	nonlinear systems control linear equations class boundary equation stability fractional output differential finite order solutions stabilization timevarying input sliding unknown	Control System
6	computing cloud smart applications internet special things issue autonomous security iot framework architecture editorial intelligent system big secure section environment	Cloud Computing
7	social information online knowledge media case role technology factors review use study perspective systematic exploring students evidence behavior software health	Knowledge Discovery

2.2 Bianchi et al. 2021 claim that their approach produces more coherent topics than previous methods. Let's test this claim by comparing the coherence of the topics produced by CTM with the topics produced by LDA. Describe your observations in 2-4 sentences.

A long standing open question is how to quantify coherence. Coherence can be measured in numerous ways, like Lau et al., 2014; Roder et al. ¨ , 2015. In this assignment, we can tell that the topics generated by CTM are more coherent simply by human observation. Although some keywords generated by the CTM models are still not coherent enough to show what is the topic. For example, in peri0d 1, topic 0, there are keywords "network", "architecture","operations"(more bottom layer) but there are also "software".
However, generally other topics are clearer than the topics modelled by LDA. For example, "Computer Vision" in period 2, includes keywords such as "image","detection","segmentation". "Computer Networks" in period 2 includes keywords such as "networks","protocal","traffic". "Deep Learning" in period 3, includes keywords such as "deep learning","classification","convolutional". "Cloud Computing" in period 3, includes keywords, such as "cloud","iot","security". We can tell that obviously topic modelled by CTM are more coherent than those modelled by LDA.

2.3 Do the two models generate similar topics? Can you discover the same temporal trends (if there are any)? Discuss in 4-6 sentences.
Yes there is certainly some overlap between the different sets of topics. For example the topic of "Computer Networks" seems to be relevant during all time periods and for both modellingn approaches. On the other hand side you can see that machine learning and deep learning keywoards only seem to become relevant after 2009.