**TOPIC MODELLING – LDA**

Pastrand doar adjective si substantive:

remove\_pos = ['ADV', 'PRON', 'PART', 'DET', 'SPACE', 'NUM', 'SYM', 'ADP', 'VERB', 'CCONJ']

stop\_words = ['paper', 'present', 'propose', 'datum', 'people', 'result', 'solution', 'case', 'order', 'base', 'ieee', 'privacy', 'policy',

'new', 'old', 'context', 'high', 'different', 'new', 'old', 'research', 'type', 'approach', 'important', 'main', 'range',

'helpful', 'large', 'difficult', 'available', 'amount', 'useful', 'importance', 'article', 'abstract', 'scale', 'copyright',

'real', 'quality', 'inconvenience', 'benefit', 'unavailable', 'term', 'condition', 'interest', 'organization', 'use',

'task', 'student', 'professor', 'teacher', 'university']

remove\_entities = ['PERSON', 'NORP', 'FAC', 'ORG', 'GPE', 'LOC', 'LANGUAGE', 'DATE', 'TIME', 'PERCENT', 'MONEY',

'QUANTITY', 'CARDINAL', 'ORDINAL']

Id = 562:

0: 0.015\*"system" + 0.007\*"application" + 0.006\*"algorithm" + 0.006\*"time" + 0.006\*"model" + 0.005\*"service" + 0.005\*"resource" + 0.005\*"performance" + 0.005\*"information" + 0.004\*"environment"

* Cuvinte nesemnificative

Pastrand doar adjective si substantive si doar bigrame:

remove\_pos = ['ADV', 'PRON', 'PART', 'DET', 'SPACE', 'NUM', 'SYM', 'ADP', 'VERB', 'CCONJ']

stop\_words = ['paper', 'present', 'propose', 'datum', 'people', 'result', 'solution', 'case', 'order', 'base', 'ieee', 'privacy', 'policy', 'new', 'old', 'context']

remove\_entities = ['PERSON', 'NORP', 'FAC', 'ORG', 'GPE', 'LOC', 'LANGUAGE', 'DATE', 'TIME', 'PERCENT', 'MONEY', 'QUANTITY', 'CARDINAL', 'ORDINAL']

lda\_model = LdaMulticore(corpus=corpus, id2word=dictionary, iterations=500, num\_topics=1, workers=4, passes=100)

Author\_id = 562:

0: 0.014\*"cloud computing" + 0.014\*"scheduling algorithm" + 0.010\*"real time" + 0.010\*"resource management" + 0.010\*"large scale" + 0.008\*"satellite image" + 0.006\*"service level" + 0.006\*"task scheduling" + 0.006\*"smart city" + 0.006\*"cloud service"

Author\_id = 534:

0: 0.011\*"large scale" + 0.011\*"mobile device" + 0.008\*"opportunistic network" + 0.005\*"real time" + 0.004\*"cloud computing" + 0.004\*"wide range" + 0.003\*"quality life" + 0.003\*"scale system" + 0.003\*"inconvenience helpful" + 0.003\*"sale account"

Author\_id = 1146:

0: 0.015\*"virtual reality" + 0.008\*"real time" + 0.007\*"virtual environment" + 0.005\*"sensory substitution" + 0.005\*"virtual space" + 0.004\*"sound localization" + 0.004\*"sound source" + 0.004\*"smith chart" + 0.003\*"virtual world" + 0.003\*"augmented reality"

Author\_id = 841:

0: 0.012\*"x ray" + 0.008\*"electron microscopy" + 0.006\*"composite material" + 0.006\*"drug delivery" + 0.005\*"ray diffraction" + 0.004\*"iron oxide" + 0.004\*"sol gel" + 0.003\*"microscopy sem" + 0.003\*"transmission electron" + 0.003\*"thin film"

Author\_id = 872:

0: 0.014\*"neural network" + 0.009\*"real time" + 0.008\*"wireless sensor" + 0.008\*"fractal dimension" + 0.007\*"unmanned aerial" + 0.006\*"sensor network" + 0.006\*"image processing" + 0.006\*"large scale" + 0.005\*"convolutional neural" + 0.005\*"control system"

Author\_id = 1284:

0: 0.011\*"natural language" + 0.008\*"e learning" + 0.006\*"textual complexity" + 0.006\*"chat conversation" + 0.005\*"social network" + 0.004\*"learning process" + 0.003\*"polyphonic model" + 0.003\*"machine learning" + 0.003\*"artificial intelligence" + 0.003\*"web page"

Cu *stop\_words = ['paper', 'present', 'propose', 'datum', 'people', 'result', 'solution', 'case', 'order', 'base', 'ieee', 'privacy', 'policy',*

*'new', 'old', 'context', 'high', 'different', 'new', 'old', 'research', 'type', 'approach', 'important', 'main', 'range',*

*'helpful', 'large', 'difficult', 'available', 'amount', 'useful', 'importance', 'article', 'abstract', 'scale', 'copyright',*

*'real', 'quality', 'inconvenience', 'benefit', 'unavailable', 'term', 'condition', 'interest', 'organization', 'use',*

*'task', 'student', 'professor', 'teacher', 'university']*

lda\_model = LdaMulticore(corpus=corpus, id2word=dictionary, iterations=**500**, num\_topics=1, workers=4, passes=**200**)

Id = 534:  
 0: 0.011\*"**mobile device**" + 0.009\*"**opportunistic network**" + 0.005\*"**cloud computing**" + 0.004\*"information product" + 0.004\*"service updated" + 0.004\*"fault tolerance" + 0.004\*"technology humanity" + 0.004\*"site agreement" + 0.004\*"profit world" + 0.004\*"account management"

Id = 562:

0: 0.017\*"**scheduling algorithm**" + 0.016\*"**cloud computing**" + 0.010\*"**resource management**" + 0.008\*"**satellite image**" + 0.007\*"**service level**" + 0.007\*"**smart city**" + 0.007\*"**cloud service**" + 0.006\*"**time series**" + 0.006\*"**fault tolerant**" + 0.006\*"**genetic algorithm**"

Id = 1146:

0: 0.016\*"**virtual reality**" + 0.007\*"**virtual environment**" + 0.005\*"**sound localization**" + 0.005\*"**sensory substitution**" + 0.005\*"**virtual space**" + 0.005\*"**assistive device**" + 0.004\*"virtual world" + 0.004\*"smith chart" + 0.003\*"fear level" + 0.003\*"sound source"

Id = 841:

0: 0.013\*"**x ray**" + 0.010\*"**electron microscopy**" + 0.007\*"**drug delivery**" + 0.006\*"**composite material**" + 0.005\*"**ray diffraction**" + 0.004\*"iron oxide" + 0.004\*"sol gel" + 0.004\*"microscopy sem" + 0.003\*"thin film" + 0.003\*"mesoporous silica"

Id = 872:

0: 0.016\*"**neural network**" + 0.009\*"**fractal dimension**" + 0.009\*"**wireless sensor**" + 0.007\*"**unmanned aerial**" + 0.006\*"**image processing**" + 0.006\*"**sensor network**" + 0.005\*"**convolutional neural**" + 0.005\*"**artificial intelligence**" + 0.005\*"**energy consumption**" + 0.005\*"**optic disc**"

Id = 1284:

0: 0.012\*"**natural language**" + 0.009\*"**e learning**" + 0.007\*"**textual complexity**" + 0.007\*"**chat conversation**" + 0.006\*"**social network**" + 0.004\*"polyphonic model" + 0.004\*"learning process" + 0.003\*"artificial intelligence" + 0.003\*"machine learning" + 0.003\*"collaborative learning"

*Adaugand trigrame:*

*remove\_pos = ['ADV', 'PRON', 'PART', 'DET', 'SPACE', 'NUM', 'SYM', 'ADP', 'VERB', 'CCONJ']*

*STOP\_WORDS = ['paper', 'present', 'propose', 'show', 'datum', 'people', 'result', 'solution', 'case', 'order',*

*'base', 'ieee', 'privacy', 'policy', 'new', 'old', 'context', 'high', 'different', 'research', 'type',*

*'approach', 'important', 'main', 'range', 'helpful', 'large', 'difficult', 'available', 'amount',*

*'useful', 'importance', 'article', 'abstract', 'scale', 'copyright', 'real', 'quality', 'demonstrate',*

*'inconvenience', 'benefit', 'unavailable', 'term', 'condition', 'interest', 'recent', 'obtain',*

*'title', 'jat', 'jats',*

*'organization', 'task', 'student', 'professor', 'teacher', 'university', 'workshop', 'study', 'text',*

*'conference']*

*remove\_entities = ['PERSON', 'NORP', 'FAC', 'GPE', 'LOC', 'LANGUAGE', 'DATE', 'TIME', 'PERCENT', 'MONEY',*

*'QUANTITY', 'CARDINAL', 'ORDINAL']*

829 - DATCU Mihai

0: 0.013\*"sar image" + 0.008\*"synthetic aperture radar sar" + 0.008\*"satellite image" + 0.005\*"earth observation" + 0.005\*"feature extraction" + 0.004\*"image content" + 0.004\*"terrasar x" + 0.004\*"land cover" + 0.004\*"image patch" + 0.003\*"remote image"

1672 - GRUMEZESCU Alexandru Mihai

0: 0.004\*"drug delivery" + 0.004\*"biomedical application" + 0.004\*"drug delivery system" + 0.003\*"wound dressing" + 0.003\*"tissue engineering" + 0.003\*"magnetite nanoparticle" + 0.003\*"essential oil" + 0.003\*"cancer therapy" + 0.003\*"medical device" + 0.003\*"food industry"

841 - ANDRONESCU Ecaterina

0: 0.005\*"x ray diffraction xrd" + 0.004\*"composite material" + 0.003\*"drug delivery" + 0.003\*"mesoporous silica" + 0.003\*"drug delivery system" + 0.003\*"transmission electron microscopy tem" + 0.003\*"antimicrobial activity" + 0.002\*"thin film" + 0.002\*"magnetite nanoparticle" + 0.002\*"electron microscopy sem"

1284 - TRAUSAN-MATU STEFAN

0: 0.007\*"natural language processing" + 0.006\*"e learning" + 0.006\*"computer collaborative learning cscl" + 0.005\*"chat conversation" + 0.005\*"social network" + 0.004\*"natural language" + 0.004\*"artificial intelligence" + 0.004\*"readerbench framework" + 0.004\*"learning process" + 0.004\*"natural language processing technique"

1225 - VOICU Gheorghe

0: 0.005\*"subject category technique methodologies" + 0.005\*"mathematical model" + 0.005\*"energy consumption" + 0.004\*"subject category property" + 0.003\*"renewable energy" + 0.003\*"hammer mill" + 0.003\*"energetic plant" + 0.003\*"mechanical property" + 0.003\*"working process" + 0.003\*"raw material"

1849 - FICAI Anton

0: 0.008\*"drug delivery system" + 0.007\*"antimicrobial activity" + 0.006\*"tissue engineering" + 0.006\*"drug delivery" + 0.006\*"composite material" + 0.006\*"staphylococcus aureus" + 0.006\*"raw material" + 0.006\*"medical device" + 0.005\*"pore size" + 0.005\*"mechanical property"

534 - DOBRE Ciprian Mihai

0: 0.008\*"mobile device" + 0.006\*"opportunistic network" + 0.004\*"cloud computing" + 0.004\*"big data" + 0.004\*"information product service updated" + 0.004\*"single sale account management" + 0.003\*"fault tolerance" + 0.003\*"end user" + 0.003\*"cross section" + 0.003\*"world technical professional technology"

733 - SEMENESCU Augustin

0: 0.012\*"sustainable development" + 0.007\*"technological process" + 0.007\*"mathematical model" + 0.006\*"heat treatment" + 0.006\*"chemical composition" + 0.006\*"welding process" + 0.006\*"stainless steel" + 0.006\*"mechanical property" + 0.006\*"composite material" + 0.005\*"electric arc furnace"

69354 - VLAD MAGDALENA

0: 0.020\*"fast ion" + 0.014\*"jet iter like wall" + 0.014\*"energy resolution" + 0.014\*"h mode" + 0.013\*"magnetic field" + 0.012\*"neutron spectrometer" + 0.012\*"joint european torus" + 0.012\*"poloidal asymmetry" + 0.010\*"joint european torus jet" + 0.010\*"density profile"

562 - POP Florin

0: 0.016\*"scheduling algorithm" + 0.016\*"cloud computing" + 0.011\*"big data" + 0.008\*"resource management" + 0.006\*"time series" + 0.006\*"genetic algorithm" + 0.006\*"smart city" + 0.006\*"web service" + 0.005\*"smart environment" + 0.005\*"satellite image"

1541 - UNGUREANU Nicoleta

0: 0.010\*"agricultural soil" + 0.008\*"wastewater treatment" + 0.007\*"heavy metal" + 0.007\*"mathematical model" + 0.006\*"subject category technique methodologies" + 0.006\*"sunflower seed" + 0.006\*"active body" + 0.006\*"biological contactor" + 0.006\*"finite element" + 0.006\*"finite element method"

1297 - PETRESCU Florian Ion

0: 0.005\*"distribution mechanism" + 0.005\*"original method" + 0.004\*"human body" + 0.004\*"point view" + 0.004\*"nuclear fusion" + 0.003\*"mechanical transmission" + 0.003\*"lockheed martin" + 0.003\*"energy source" + 0.003\*"mechanical system" + 0.003\*"anthropomorphic robot"

38845 - STASTNY PETER

0: 0.031\*"t cell" + 0.016\*"b cell" + 0.015\*"hla class" + 0.015\*"endothelial cell" + 0.015\*"mica antigen" + 0.014\*"hla dr" + 0.014\*"amino acid" + 0.014\*"mental health" + 0.014\*"hla b" + 0.012\*"immune response"

1047 - BIRIS SORIN STEFAN

0: 0.012\*"subject category miscellaneous" + 0.008\*"subject category technique methodologies" + 0.008\*"energy efficiency" + 0.007\*"agro pellet" + 0.007\*"plough body ante moldboard" + 0.007\*"digital hydraulic" + 0.007\*"heavy metal" + 0.006\*"finite element method" + 0.006\*"agricultural soil" + 0.006\*"methodologies equipment"

584 - DASCALU Mihai

0: 0.012\*"e learning" + 0.010\*"cellular automata" + 0.009\*"state art" + 0.009\*"fake news" + 0.008\*"bert model" + 0.007\*"language model" + 0.007\*"natural language processing nlp" + 0.006\*"neural network" + 0.005\*"project management" + 0.005\*"social network"

872 - POPESCU Dan

0: 0.012\*"neural network" + 0.006\*"fractal dimension" + 0.006\*"sensor network" + 0.005\*"image processing" + 0.005\*"wireless sensor network" + 0.004\*"unmanned aerial vehicle uav" + 0.004\*"texture classification" + 0.004\*"energy consumption" + 0.004\*"point view" + 0.004\*"monitoring system"

1292 - RADU Gabriel Lucian

0: 0.007\*"antioxidant activity" + 0.006\*"detection limit" + 0.005\*"aqueous extract" + 0.004\*"limit detection" + 0.004\*"ascorbic acid" + 0.003\*"free radical" + 0.003\*"inhibitory activity" + 0.003\*"rosmarinic acid" + 0.003\*"capillary electrophoresis" + 0.003\*"anti inflammatory"

1246 - IOVU Horia

0: 0.006\*"graphene oxide" + 0.006\*"mechanical property" + 0.005\*"polymer matrix" + 0.004\*"thermal stability" + 0.004\*"electron microscopy" + 0.004\*"x ray photoelectron spectroscopy" + 0.004\*"hybrid material" + 0.003\*"molecular weight" + 0.003\*"electron microscopy sem" + 0.003\*"ft ir"

68995 - Meghea Aurelia

0: 0.011\*"antioxidant activity" + 0.006\*"vegetable oil" + 0.006\*"free radical" + 0.005\*"sol gel" + 0.004\*"physical chemical" + 0.004\*"lipid nanoparticle" + 0.004\*"thin film" + 0.004\*"lipid nanocarrier" + 0.004\*"uv vis" + 0.004\*"heavy metal"

1146 - MOLDOVEANU ALIN - DRAGOS - BOGDAN

0: 0.012\*"virtual reality" + 0.007\*"virtual environment" + 0.007\*"smith chart" + 0.005\*"sound vision" + 0.004\*"human body" + 0.004\*"virtual world" + 0.004\*"fear level" + 0.004\*"sensory substitution" + 0.004\*"human health" + 0.003\*"augmented reality"