VILNIUS TECH

STUDIJŲ DALYKO (MODULIO) (SD(M)) KORTELĖ PIRMOSIOS IR ANTROSIOS PAKOPŲ BEI VIENTISOSIOMS STUDIJOMS

Matematinio modeliavimo katedra

COURSE CARD FOR THE FIRST AND SECOND CYCLE AND INTEGRATED STUDIES

Department of Mathematical Modelling

SD(M) pavadinimas	***Course title***
Mašininis mokymasis	***Machine Learning***

SD(M) priklausomybė studijų pakopai

Course subjection to study level

Studijos: *Studies:*	B – Pirmosios pakopos First cycle

SD(M) priklausomybė studijų programai

Course subjection to programme

SD(M) priklausomybė dalykų grupei * *Course subjection to group*	1 – studijų dalyko Course
SD(M) priklausomybė programos daliai ** *Course subjection to part of the programme*	C – Specializacijos dalykų dalis Part of Specialization Subjects
Struktūrinė SD priklausomybė *** *Course structural subjection*	K – katedros Department

SD(M) priklausomybė studijų krypčių ir krypčių grupei

The list of study fields and groups of fields

Studijų krypčių grupės kodas *Code of the group of study fields*	Studijų krypties kodas *Code of the study field*
A	A01

***) Grupė:** *) 1 - studijų dalyko; 2 - praktikos; 3 - baigiamojo darbo ar projekto; 4 - baigiamojo egzamino; 5 - tiriamojo darbo; 6 - profesinio testavimo; 7 - kitas.
**) A - Bendrųjų universitetinių studijų; B - Studijų krypties; C - Specializacijos.
***) U - universiteto; F - fakulteto; K - katedros.

***) Group:** *) 1 - *Course*; 2 - *Practice*; 3 - *Final Work or Project*; 4 - *Final Examination*; 5 - *Research Work*; 6 - *Professional Testing*; 7 - *Other*.
**) A - *General*; B - *Field*; C - *Specialization*.
***) U - *University*; F - *Faculty*; K - *Department*.

SD(M) kodas

Course number

Fakultetas *Faculty*		Katedra *Department*		Pakopa * *Study cycle*	Modulio Nr. *Number*
F	M	M	M	B	21501

SD(M) kreditai

Course volume in credits

Iš viso: *Total:*	Iš jų: KD, KS, KP, PR *There out:*
3	0

SD(M) Atsiskaitymo forma

Course assessment

Į, E1, E2, E, BE, BD, TD, A	KD, KS, KP, PR
E	-

*) B - pirmoji pakopa; A - vientisosios studijos; M - antroji pakopa.

*) B - *first cycle studies*; A - *integrated studies*; M - *second cycle studies*.

SD(M) valandų paskirstymas pagal studijų formas ir būdus

Distribution of course hours by study forms and ways

Studijų forma *Study form*	Valandos *Hours*								Kontaktinių *Contact*
	Kodas *Code*	Studijų būdas * *Study way*	Paskaitoms *Lectures*	Lab. darbams *Laborotory works*	Pratyboms *Practical works*	Konsultacijoms *Consultation*	Sav. darbui *Independent work*	Iš viso *Total*	
Nuolatinės studijos *Full-time studies*	NL	S	15	15	0	2	48	80	32

*) Studijų būdas: S - semestrais; M - moduliais; C - ciklais; T - nuotolinis; NI - neakivaizdinis intensyvusis.

*) Study process forms: S - *semesters*; M - *modules*; C - *periods*; T - *distance*; NI - *part-time*.

SD(M) ANOTACIJA

Mašininis mokymasis šiandien yra neatsiejama daugelio komercinių programų ir tyrimų projektų dalis pradedant medicinine diagnostika ir gydymu, baigiant draugų radimu socialiniuose tinkluose. Daugeliu atveju ši metodika palengvina įvairių organizacijų darbą. Šiame studijų dalyke parodysime, kaip lengvai ir su teoriniu pagrindimu galima susikurti mašininio mokymo sprendimus ir kaip tai padaryti geriausiai.

ANNOTATION OF COURSE

Machine learning is an integral part of many commercial applications and research projects today, in areas ranging from medical diagnosis and treatment to finding your friends on social networks. In many cases this methodology facilitates the work of various organizations. In this course, we want to show you how easy it can be to build machine learning solutions with theoretical background yourself and how to do best it.

SD(M) TIKSLAS

supažindinti su:
- pagrindiniais mašininio mokymosi tipais ir metodais realiems uždaviniams spręsti;
- metodų teoriniais pagrindais, jų taikymo galimybėmis (privalumais ir trūkimais);
- Python (arba R) programinio paketo galimybėmis šiems uždaviniams spręsti.

AIM OF COURSE

Is to introduce to:
-the main types ant methods of machine learning to solve real problems;
-theoretical foundation of methods, their application possibilities (adventages and disadventages);
-Python (or R) software capabilities to solve these challenges.

Pagrindinė literatūra (ne daugiau kaip 5 šaltiniai):

Main references (not more than 5 references)

Eil. Nr. *No.*	Leidinio autoriai ir pavadinimas (elektroninių leidinių ir žiniatinklio adreso) *Authors and title (site address in case of e-publication)*
1.	Trevor Hastie, Robert Tibshirani, Jerome Friedman. The elements of statistical learning : data mining, inference, and prediction. Springer, 2017. Neribota prieiga oficialiu adresu: https://web.stanford.edu/~hastie/Papers/ESLII.pdf
2.	Ian Goodfellow and Yoshua Bengio and Aaron Courville. Deap Learning, 2016, MIT press http://www.deeplearningbook.org
3.	Witten, I.H., E. Frank. Data Mining: Practical Machine Learning Tools and Techniques, 2011. Prieiga per VPN: https://www.sciencedirect.com/book/9780123748560/data-mining-practical-machine-learning-tools-and-techniques
4.	Sutton, R. S., A. G. Barto. Reinforcement Learning: An Introduction. Neribota prieiga oficialiu adresu: https://web.stanford.edu/class/psych209/Readings/SuttonBartoIPRLBook2ndEd.pdf

*) Kortelės pildymo metu

*) *At the form filling moment*

Papildoma literatūra (ne daugiau kaip 10 šaltinių):

Additional references (not more than 10 references)

Eil. Nr. *No.*	Leidinio autoriai ir pavadinimas (elektroninių leidinių ir žiniatinklio adreso) *Authors and title (site address in case of e-publication)*
1.	Theodoridis S., K. Koutroumbas. Pattern Recognition Prentice-Hall, 2006 neribota prieiga internete per VGTU ScienceDirec
2.	Russel, S., P. Norvig Artificial Intelligence: A Modern Approach (2nd ed.) Prentice-Hall, 2002
3.	Sutton, R. S., A. G. Barto. Reinforcement Learning: An Introduction The MIT Press, 2017 http://incompleteideas.net/book/bookdraft2017nov5.pdf

*) Kortelės pildymo metu

*) A*t the form filling moment*

Reikalingi IT resursai * (nurodyti 1-3 alternatyvas, pageidautina, kad bent 1 būtų nemokama)

Required IT Resources

Eil. Nr. *No.*	Programinės įrangos pavadinimas, gamintojas *Name of the sofware, manufacturer*	Licencijos tipas (pagal įsigijimo būdą) *License type*
1	Python (Anaconda Navigator)	Nemokama Unpaid
2	R; R studio	Nemokama Unpaid

*) Pildoma, jei tokie resursai reikalingi. Stulpelyje Licencijos tipas pasirenkamas iš sąrašo:

Mokama, akademinė
Mokama, komercinė
Nemokama
*) *Should be completed if such reassures are needed. License type - select from the list:*
Paid, academic
Paid, commercial

Savarankiško darbo turinys

Content of individual work

Užduoties pavadinimas *Assignment title*	Sav. darbo apimtis vienai užduočiai *Amount of hours of independent work for a single task*						Užduočių skaičius *Number of tasks*					Iš viso valandų *Total hours*					Įvertinimo dalis % *Part of Evaluation %*				
	Rekomenduojamos val. *Recommended hours*	Skirta val. Separated hours					NL (T)	NL (S)	NL (Sav.)	I(S)	I(T)	NL(T)	NL(S)	NL (Sav.)	I(S)	I(T)	NL(T)	NL(S)	NL (Sav.)	I(S)	I(T)
		NL(T)	NL(S)	NL (Sav.)	I(S)	I(T)															
Kolokviumas Intermediate examination	8-27		18					1					18					20			
Laboratorinis darbas Laboratory work	2-12		5					3					15					30			
Pasirengimas atsiskaitymui Preparation for evaluation	10-60		15					1					15								
Iš viso: *Total:*													48								

*) Papildomas laukas pildomas tik tada, kada taikomas SD(M) kortelėje nenurodytas studijų būdas: M - moduliais; C - ciklais; T - nuotolinis

*) M*ust be used in case study way does not fall into standard category*: M - *modules*; C - *periods*; T - *distance*

Savarankiško darbo grafikas

Individual work schedule

Užduoties tipas ***Task type***	Užduoties pateikimo(*) ir atsiskaitymo(+) savaitė *Week of Assignment setting (*) and assessment(+)*																				
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Nuolatinės studijos (S) ***Full-time studies***																					
Kolokviumas *Intermediate examination*	* +				1			1													
Laboratorinis darbas *Laboratory work*	* +					1	1	2	2		3	3									

Laboratorinių darbų sąrašas

List of the Course laboratory work

Temos pavadinimas *Topic title*	Valandų skaičius *Number of hours*				
	NL(T)	NL(S)	NL(Sav.)	I(S)	I(T)
1. Python Pandas (R studio). Duomenų grupavimas, išskyrimas, statistinių charakteristikų skaičiavimas ir vaizdavimas *Python Pandas. Data extraction, grouping and representation of statistical characteristics.*		3			
2. Duomenų skaidymas. Modelio permokymas ir nepakankamas apmokymas tiesinei regresijai. Gradientinio nusileidimo metodas. *Data splitting. Bias, variance, underfitting and overfitting for linear regression. Gradient descent.*		2			
3. Klasifikavimo uždavinių realizavimas. Logistinė regresija. *Data classification. Logistic regression*		3			
4. Sprendimų medžio konstravimas klasifikavomi uždaviniui spręsti. Sprendimo medžio permokymas ir jo šakų kapojimas. Atsitiktinis miškas *Decission tree classifier. Decision tree overfitting and pruning. Random forest*		3			
5. Atraminių vektorių mašina. *Support vector machines*		2			
6. Pagrindinių komponenčių analizė *Principal component analysis*		2			
Iš viso: *Total:*		**15**			

Paskaitų temų sąrašas
List of the Course lecture topics

Temos pavadinimas *Topic title*	Valandų skaičius *Number of hours*				
	NL(T)	NL(S)	NL(Sav.)	I(S)	I(T)
1. Įvadas į mašininį mokymąsi. Mašininio mokymosi pagrindiniai tipai (mokymasis su mokytoju, mokymas be mokytojo). Duomenų skaidymo strategijos. Modelio variacija ir poslinkis, persimokymas ir nepakankamas apsimokymas tiesinei regresijai. *Introduction to machine learning. Basic types of machine learning (supervised and unsupervised learning). Data splitting. Bias, variance, underfitting and overfitting for linear regression.*		4			
2. Duomenų klasifikavimas. Paklaidų analizė. Logistinė regresija. Bajeso klasifikatorius. Modelių reguliarizacija. *Data classification. Error analysis. Logistic regression. Bayesian Classifier. Regularization.*		3			
3. Sprendimų medžių klasifikatorius. Informacijos matavimo matai. Sprendimų medžio šakų kapojimas. Regresiniai medžiai. *Decission tree classifier. Quantities of information. Decision tree pruning. Decision tree regressor.*		2			
4. Atraminių vektorių mašina. Branduolio funkcija. *Support vector machines. Kernel functions.*		2			
5. Metodų ansambliai. *Ensembles.*		2			
6. Dimensijos mažinimo uždaviniai. Principinių komponenčių analizė. *Dimensionality reduction. Principal component analysis.*		2			
Iš viso: *Total:*		**15**			

*) Papildomas laukas pildomas tik tada, kada taikomas SD(M) kortelėje nenurodytas studijų būdas: M - moduliais; C - ciklais; T - nuotolinis

*) M*ust be used in case study way does not fall into standard category*: M - *modules*; C - *periods*; T - *distance*

Fundamentinių mokslų fakulteto *Moderniųjų technologijų matematikos (612G12001)* 2016-07-01 programos studijų rezultatų sąsajos su SDM rezultatais bei studijų ir studentų pasiekimų vertinimo metodais

Links of the Mathematics of Modern Technologies (612G12001) of the Faculty of Fundamental Sciences with the course unit and evaluation methods of students achievements

Programos studijų rezultatai *Study programme outcomes*	SD(M) rezultatai *Course results*	Studijų metodai *Methods of studies*	Studento pasiekimų vertinimo metodai *Evaluation methods of student archievements*	Studentų pasiekimų vertinimo kriterijai pagal lygmenis *Assessments criteria of students achievements by Assessment levels*
GT1. Gebės įžvelgti tarpdalykinius ryšius ir į problemą pažvelgti kaip į visumą, savarankiškai analizuos problemas, nagrinės ir vertins matematinių modelių tinkamumą, savybes, supras matematinių įrodymų būtinumą, griežtumą. GT1. Will be able to spot interdisciplinary relationships and look at the problem as a whole, analyze the problems independently, analyze and assess the properties of mathematical models, as well as the overall suitability of those models, and will understand the necessity and rigor of mathematical proof.	Gebės savarankiškai išanalizuoti problemą ir jos sprendimui pritaikyti tinkamiausią ekonometrinį modelį. Students will be able to independently analyze the problem and to choose the most appropriate econometric model for its solution.	Teorinės paskaitos, laboratoriniai darbai, savarankiškos dalyko studijos, konsultacijos. Theoretical lectures, exercises, laboratory works, independent study of subjects, consultations	Laboratorinių darbų vertinimas, kolokviumas, egzaminas. Laboratory work assessment, colloquium and exam.	Slenkstinis: studentas žino pagrindinius ekonometrijos principus ir dėsnius, moka įgytas žinias taikyti sudarydamas paprasčiausius modelius programinio paketo pagalba. Tipinis: studentas žino pagrindinius ekonometrijos principus ir dėsnius, geba įgytas žinias taikyti atlikdamas praktinių užduočiou analizę, skaičiavimams taiko kompiuterines programas, tačiau daro neesminių klaidų pateikdamas ir interpretuodamas gautus rezultatus. Puikus: studentas puikiai žino visą teorinį kursą, gali įrodyti svarbiausias teoremas, geba savarankiškai

				sudaryti naujų problemų. Threshold: The student knows basic concepts and laws of Econometrics and can apply them to create simple models using econometric analysis software package . Typical: the student knows the basic concepts and laws of Econometrics, is able to apply the acquired knowledge through analysis of practical tasks, applies computational software for calculations, but makes minor mistakes when presenting and interpreting the obtained results. Excellent: the student knows perfectly the whole theoretical course, is able to independently create sophisticated mathematical models of physical phenomena, analyze and evaluate them by using computer programs, presents and interprets the obtained results.
GT2. Gebės surinkti duomenis ir juos analizuoti, taikys įvairius jų peržiūros algoritmus, siūlys naujus technologinius sprendimus, kurs naujus matematinius modelius ir tobulins jau sukurtus. GT2. Will be able to collect and analyze data, apply their various review algorithms, offer new technological solutions, develop new mathematical models and improve the existing ones.	Gebės pritaikyti duomenų atrankos principus, susisteminti, analizuoti surinktus duomenis. Students recieve ability to apply the data selection principles, organize, represent graphically and analyze the collected data.	Teorinės paskaitos, laboratoriniai darbai, savarankiškos dalyko studijos, konsultacijos. Theoretical lectures, exercises, laboratory works, independent study of subjects, consultations.	Laboratorinių darbų vertinimas, kolokviumas, egzaminas Laboratory work assessment, colloquium and exam.	Slenkstinis: studentas žino pagrindinius ekonometrijos principus ir dėsnius, moka įgytas žinias taikyti sudarydamas paprasčiausius modelius programinio paketo pagalba. Tipinis: studentas žino pagrindinius ekonometrijos principus ir dėsnius, geba įgytas žinias taikyti atlikdamas praktinių užduočiou analizę, skaičiavimams taiko kompiuterines programas, tačiau daro neesminių klaidų pateikdamas ir interpretuodamas gautus rezultatus. Puikus: studentas puikiai žino visą teorinį kursą, gali įrodyti svarbiausias teoremas,

				geba savarankiškai sudaryti naujų problemų. Threshold: The student knows basic concepts and laws of Econometrics and can apply them to create simple models using econometric analysis software package . Typical: the student knows the basic concepts and laws of Econometrics, is able to apply the acquired knowledge through analysis of practical tasks, applies computational software for calculations, but makes minor mistakes when presenting and interpreting the obtained results. Excellent: the student knows perfectly the whole theoretical course, is able to independently create sophisticated mathematical models of physical phenomena, analyze and evaluate them by using computer programs, presents and interprets the obtained results.

SD(M) sudarytojas (-ai) (parašas, vardas ir pavardė)
Course compiled by *(full name, signature)*
Gerda Jankevičiūtė

Katedros vedėjas (parašas, vardas ir pavardė)
Head of Department *(full name, signature)*
Raimondas Čiegis

SD(M) atestuojamas *The Course is certified*				
SD(M), skirtas studijų programai: *The Course for the programme of studies***:**	**Moderniųjų technologijų matematika**			
SD(M) atestacija galioja: *Course certification is valid:*	nuo *from*		iki *till*	

SD(M) atestavo *the Course certified by*	Fundamentinių mokslų fakulteto studijų komitetas		
Fakulteto studijų komiteto pirmininkas (vardas ir pavardė, parašas) *Chairman of the Studies committee (full name, signature)*		Data *Date*	