BIOINFORMATIKA – INDIVIDUALNO RAZISKOVALNI PREDMETI

UČNI NAČRT PREDMETA/COURSE SYLLABUS

Predmet:	Bioinformacijski algoritmi
Course title:	Bioinformatics Algorithms

Študijski programi in stopnja	Študijska smer	Letnik	Semestri
Bioznanosti, tretja stopnja, doktorski	bioinformatika		Celoletni

Univerzitetna koda predmeta/University course code: 3765

Predavanja	Seminar	Vaje	Klinične vaje	Druge oblike študija	Samostojno delo	ECTS
10	0	20	0	0	95	5

Nosilec predmeta/Lecturer:	Tomaž Curk

Izvajalci predavanj:	
Izvajalci seminarjev:	Tomaž Curk, Blaž Zupan
Izvajalci vaj:	
Izvajalci kliničnih vaj:	
Izvajalci drugih oblik:	
Izvajalci praktičnega usposabljanja:	

Vrsta predmeta/Course type:	individualno raziskovalni/individual research course

Jeziki/Languages:	Predavanja/Lectures:	Angleščina, Slovenščina
	Vaje/Tutorial:	Angleščina, Slovenščina

Pogoji za vključitev v delo oz. za opravljanje študijskih obveznosti:	Prerequisites:
Osnovna znanja programiranja, verjetnosti in statistike.	Basics of computer programming in any language, basic knowledge of probability and statistics.

Vsebina:	Content (Syllabus outline):
- Algoritmi za analizo zaporedij, iskanje podzaporedij, iskanje motivov. - Določanje zaporedja genomov, algoritmi na grafih. - Primerjava zaporedij, dinamično programiranje. - Algoritmi za filogenetsko analizo. - Skriti markovski modeli, analiza strukture genoma. - Analiza genskih izrazov, razvrščanje v skupine, klasifikacija, analiza obogatenosti genskih skupin. - Rekonstrukcija in analiza genskih mrež. - Vizualizacija podatkov.	- Sequence analysis, search for subsequences, motif search. - Genome assembly, graph algorithms. - Comparison of biological sequences, dynamic programming. - Phylogeny algorithms. - Hidden Markov Models and gene prediction. - Gene expression analysis, clustering and supervised data mining, enrichment analysis. - Gene network reconstruction and analysis.

	- Data vizualization.

Temeljna literatura in viri/Readings:

- Durbin R, Eddy SR, Krogh A, Mitchison G (1998) Biological sequence analysis: probabilistic models of proteins and nucleic acids, Cambridge University Press.
- Jones NC, Pevzner PA (2004) An introduction to bioinformatics algorithms, The MIT Press.

Ostalo: revijalni članki s področja, tekoča periodika in druga učna gradiva.

Cilji in kompetence:	**Objectives and competences:**
Študentje se bodo pri predmetu naučili implementirati vrsto algoritmov, ki jih lahko uporabimo na področju bioinformatike in sistemske biologije. Znali bodo razbrati, na kater tip biološki vprašanj lahko odgovorimo z razvojem in uporabo računskih pristopov.	Students completing the course should be able to implement a variety of bioinformatics and systems biology algorithms, and learn which type of biological questions can be answered by means of computational approaches.

Predvideni študijski rezultati:	**Intended learning outcomes:**
Študentje se bodo seznanili z glavnimi razredi algoritmov, ki so uporabljajo na področju bioinformatike in lahko z njimi analiziramo zaporedja, grafe in podatke o meritvah iz molekularne biologije. Na praktičnih primerih analize velike množice podatkov bodo spoznali probleme pri razvoju teh algoritmov, ki so vezani na hitrost izvajanja in uporabo spomina. Izpopolnili bodo svoje predznanje programiranja in v praksi uporabili predznanja s področja verjetnosti in statisike.	Students will become familiar with main classess of computational approaches and algorithms in bioinformatics. The algorithms that they will design in a class are those from sequence and graph analysis and analysis of data coming from experimental measurements in molecular biology. In practical cases of analysis of large data sets they will need to cope with problems of computational efficiency and limited data storage (computer memory). They will advance their knowledge of programming, and use thier previously developed skills in probability and statistics in practical problems from systems biology.

Metode poučevanja in učenja:	**Learning and teaching methods:**
Praktične vaje, domače naloge, seminar, konzultacije. Reševanje problemov na učnih spletnih straneh http://rosalind.info in http://stepic.org.	Workshops, homeworks, consultations, seminar. Solving problems on learning portals such as http://rosalind.info and http://stepic.org.

Načini ocenjevanja:	**Delež/Weight**	**Assessment:**
Domače naloge	20,00 %	Homeworks
seminar	80,00 %	seminar

Reference nosilca/Lecturer's references:

Blaž Zupan:
1. Toplak M, Močnik R, Polajnar M, Bosnić Z, Carlsson L, Hasselgren C, Demšar J, Boyer S, Zupan B, Stålring J (2014) Assessment of Machine Learning Reliability Methods for Quantifying the Applicability Domain of QSAR Regression Models, J Chem Inf Model 54(2):431-41.
2. Žitnik M, Janjić V, Larminie C, Zupan B, Pržulj N (2013) Discovering disease-disease associations by fusing systems-level molecular data, Scientific Reports, 13:3202. [COBISS.SI-ID 10253396]
3. Demsar J, Curk T, Erjavec A, Gorup C, Hocevar T, Milutinovic M, Mozina M, Polajnar M, Toplak M, Staric A, Stajdohar M, Umek L, Zagar L, Zbontar J, Zitnik M, Zupan B (2013) Orange: data mining toolbox in Python, Journal of Machine Learning Research, 14:2349-2353. [COBISS.SI-ID 10118740]

4. Nasser W, Santhanam B, Miranda ER, Parikh A, Juneja K, Rot G, Dinh C, Chen R, Zupan B, Shaulsky G, Kuspa A (2013) Bacterial discrimination by dictyostelid amoebae reveals the complexity of ancient interspecies interactions, Current Biology, 23(10):862-872. [COBISS.SI-ID 9921108]
5. Miranda ER, Zhuchenko O, Toplak M, Santhanam B, Zupan B, Kuspa A, Shaulsky G (2013) ABC transporters in Dictyostelium discoideum development, PLoS One, 8 (8). e70040. [COBISS.SI-ID 10035284]
6. Mendes M, Franco-Duarte R, Umek L, Fonseca E, Drumonde-Neves J, Dequin S, Zupan B, Schuller D (2013) Computational models for prediction of yeast strain potential for winemaking from phenotypic profiles, PLoS One, 8 (7). e66523. [COBISS.SI-ID 10005844]

Tomaž Curk:
1. Yoichiro Sugimoto, Julian König, Shobbir Hussain, Blaž Zupan, Tomaž Curk, Michaela Frye, Jernej Ule (2012) Analysis of CLIP and iCLIP methods for nucleotide-resolution studies of protein-RNA interactions, Genome Biology 13(8): 1-33. [COBISS.SI-ID 9319764]
2. Tomaž Curk, Gregor Rot, Blaž Zupan (2011) SNPsyn : detection and exploration of SNP-SNP interactions, Nucleic Acids Research 39(2): 444-449. [COBISS.SI-ID 8352596]
3. James R. Tollervey*, Tomaž Curk*, Boris Rogelj*, Michael Briese, Matteo Cereda, Melis Kayikci, Julian König, Tibor Hortobágyi, Agnes L. Nishimura, Vera Župunski, Rickie Patani, Siddharthan Chandran, Gregor Rot, Blaž Zupan, Christopher E. Shaw, Jernej Ule (2011) Characterizing the RNA targets and position - dependent splicing regulation by TDP-43, Nature Neuroscience 14(4): 452-459. [COBISS.SI-ID 8278100]
4. Julian König, Kathi Zarnack, Gregor Rot, Tomaž Curk, Melis Kayikci, Blaž Zupan, Daniel J. Turner, Nicholas M. Luscombe, Jernej Ule (2010) ICLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution, Nature Structural and Molecular Biology 17(7): 909-916. [COBISS.SI-ID 7800916]
5. Tomaž Curk, Uroš Petrovič, Gad Shaulsky, Blaž Zupan (2009) Rule-based clustering for gene promoter structure discovery, Methods of Information in Medicine 48(3): 229-235. [COBISS.SI-ID 22605095]
6. Petra Fey, Pascale Gaudet, Tomaž Curk, Blaž Zupan, Eric M. Just, Basu Siddhartha, Sohel N. Merchant, Yulia A. Bushmanova, Gad Shaulsky, Warren A. Kibbe, Rex L. Chisholm (2009) DictyBase - a Dictyostelium bioinformatics resource update, Nucleic Acids Research 37: 515-519. [COBISS.SI-ID 6916180]