FACULTY OF CHEMISTRY

SUBJECT CARD

Name of subject in Polish   **Bioinformatyka**
Name of subject in English   **Bioinformatics**
Main field of study (if applicable): **Biosciences**
Specialization (if applicable):   **Bioinformatics**
Profile: **academic**
Level and form of studies: **2nd level, full-time**
Kind of subject: **obligatory**
Subject code W03BSS-SM2004W, W03BSS-SM2004L
Group of courses NO

	Lecture	Classes	Laboratory	Project	Seminar
Number of hours of organized classes in University (ZZU)	30		30		
Number of hours of total student workload (CNPS)	75		50		
Form of crediting (Examination / crediting with grade)	Examination		crediting with grade		
For group of courses mark (X) final course					
Number of ECTS points	3		2		
including number of ECTS points for practical classes (P)			2		
including number of ECTS points corresponding to classes that require direct participation of lecturers and other academics (BU)	1,3		1,4		

*delete as not necessary

PREREQUISITES RELATING TO KNOWLEDGE, SKILLS AND OTHER COMPETENCES
1. Knowledge of basics concepts and vocabulary of biology, biochemistry and genetics
2. Ability to use computer, Internet and command line (shell) interface
3. Fluent command of English language
\

SUBJECT OBJECTIVES
C1 Teaching the retrieval of specific information from databases of biosequences, genes, genomes, structures, protein families and other biochemistry and medicine related databases.
C2 Understanding of various sequence similarity measures ant their interpretation required to perform comparative analysis of multiple sequences.
C3 Ability to search for homologous sequences, creation and use of sequence similarity profiles ant to analyze relations between sequences.
C4 Ability to build and evaluate protein models using contemporary structure prediction methods.
C5 Ability to automate typical bioinformatics analyzes and searches using self programmed scripts using specialized libraries.

SUBJECT EDUCATIONAL EFFECTS

Relating to knowledge:

PEU_W01 Knowledge on the availability and scope of information available in databases of sequences, sequence families, genes, genomes, biochemical and medicinal databases.

PEU_W02 Knowledge on the theoretical foundations of methods employed to compare sequences and analyze their features and function, necessary to evaluate the statistical significance of the results.

PEU_W03 Knowledge on the theoretical foundations of methods used for comparative analyzes, their advantages, disadvantages and scope

PEU_W04 Knowledge of the contemporary methods of structure prediction

Relating to skills:

PEU_U01 Ability to construct complex queries against the databases of sequence, gene, genome, biochemical and medicinal information to retrieve the specific data or sequences

PEU_U02 Ability to search sequence databases for similar sequences, including searches with similarity profiles, to identify homologs

PEU_U03 Ability to calculate, edit and employ multiple sequence alignments to recognize features, functions, structure and phylogenesis and other comparative analyzes of biosequence database

PEU_U04 Ability to automate common bioinformatics tasks and analyses using scripting programming languages and specialized libraries

PROGRAMME CONTENT

Lecture		Number of hours
Lec 1	Bioinformatics databases	2
Lec 2	Efficient use of databases: annotations, organization of information, complex queries against specified record fields.	2
Lec 3	Similarity and homology. Evaluation and interpretation of sequence similarity and of its statistical significance.	2
Lec 4	Methods of sequence alignments. Theoretical basis of methods of similarity-based database searches.	2
Lec 5	Approaches to multiple sequence alignment problem.	2
Lec 6	Similarity profiles as representation of sequence similarity and features. Families of sequences and databases of families of homologs. Database queries using sumilarity profiles.	2
Lec 7	Introduction to Bayesian statistics and interpretation of information encoded in bilogical sequences	2
Lec 8	Hidden Markov Models, machine learning methods and stochastic optimization approaches – applications in bioinformatics.	2
Lec 9	Theoretical models and calculation of evolutionary distances.	2
Lec 10	Methods of molecular phylogenetic analysis: inferring relations and mutation history among related sequences	2
Lec 11	Structure prediction methods, model evaluation and optimization	2
Lec 12	Automation of common bioinformatics tasks and analyses: bioinformatics programming APIs and libraries	2
Lec 13	Automation of sequence analysis, structure prediction and other tasks	2

Lec 14	Contemporary research, analytic and diagnostic techniques.	4
	Total hours	30

Laboratory		Number of hours
Lab 1	Introduction to course topics, organization of the lab, required software. Introduction to NCBI databases and the Entrez System	2
Lab 2	Complex queries. Available sequence, gene, genome and secondary databases. Different search engines.	2
Lab 3	Uniprot KB, Protein Data Bank, Brenda Enzymes and a selection of other databases related to biochemistry and medicine.	2
Lab 4	Individual task #1	2
Lab 5	Searching for similar sequences using BLAST variants. Interpretation of results.	2
Lab 6	Searching for remote homology using similarity profiles	2
Lab 7	Individual task #2	2
Lab 8	Calculation, analysis, verification and visualization of multiple sequence alignments	2
Lab 9	Use of Python scripting language and Biopython library for automation of database queries and calculations	2
Lab 10	Phylogenetic analysis	2
Lab 11	Statistical evaluation of results using bootstrap analysis	2
Lab 12	Individual task #3	2
Lab 13	Protein structure prediction based on templates.	2
Lab 14	Ab initio protein structure prediction. Evaluation of models.	2
Lab 15	Individual task #4	2
	Total hours	30

TEACHING TOOLS USED

N1. Lecture with multimedia presentation
N2. Problem solving
N3. Use of specialized software
N4. Preparation of reports of individial tasks, with analysis of results

EVALUATION OF SUBJECT LEARNING OUTCOMES ACHIEVEMENT

Evaluation (F – forming during semester), P – concluding (at semester end)	Learning outcomes code	Way of evaluating learning outcomes achievement
F1_Lab	PEU_U01	Individual task #1
F2_Lab	PEU_U02	Individual task #2
F3_Lab	PEU_U03, PEU_U05	Individual task #3
F4_Lab	PEU_U04, PEU_U05	Individual task #4
P_Lecture: grade based on exam score	Score Grade	
P_Lab: grade based on total score F1_Lab+F2_Lab+F3_Lab+F4_Lab	50-59,99% 3,0 60-69,99% 3,5 70-79,99% 4,0 80-89,99% 4,5 90-100% 5,0	

PRIMARY AND SECONDARY LITERATURE

PRIMARY LITERATURE:

[1] S.Q. Ye, Bioinformatics. A practical approach, Chapman & Hall/CRC, 2008
[2] I. Eidhammer, I. Johanssen, W.R. Taylor, Protein Bioinformatics - an algorythmic approach to sequence and structure analysis, Wiley, 2004
[3] P.E. Bourne & H. Weissig (ed.), Structural Bioinformatics, Wiley, 2003
[4] A.D. Baxevanis, B.F.F. Oullette, Bioinformatics, Wiley, 2001

SECONDARY LITERATURE:

[1] The National Center for Biotechnology Information (NCBI) Handbook:
     https://www.ncbi.nlm.nih.gov/books/NBK21101/
[2] Documentation of used WWW services (available online)
[3] http://www.ncbi.nlm.nih.gov/guide/training-tutorials/

SUBJECT SUPERVISOR (NAME AND SURNAME, E-MAIL ADDRESS)

Paweł Kędzierski, Pawel.Kedzierski@pwr.edu.pl