

| FACULTY OF CHEMISTRY | | | | | |
|--|----------------------|---------|----------------------|---------|---------|
| SUBJECT CARD | | | | | |
| Name of subject in Polish | Bioinformatyka | | | | |
| Name of subject in English | Bioinformatics | | | | |
| Main field of study (if applicable): | Biotechnology | | | | |
| Specialization (if applicable): | Bioinformatics | | | | |
| Profile: | academic | | | | |
| Level and form of studies: | 2nd level, full-time | | | | |
| Kind of subject: | obligatory | | | | |
| Subject code: | BTC024011 | | | | |
| 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) | 120 | | 60 | | |
| Form of crediting | Examination | | crediting with grade | | |
| For group of courses mark final course with (X) | | | | | |
| Number of ECTS points | 4 | | 2 | | |
| including number of ECTS points for practical (P) classes | | | 2 | | |
| including number of ECTS points for direct teacher-student contact (BK) classes | 1 | | 1 | | |
| 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 and their interpretation required to perform comparative analysis of multiple sequences. C3 Ability to search for homologous sequences, creation and use of sequence similarity profiles and 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 LEARNING OUTCOMES 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. PEK_W03 Knowledge on the theoretical foundations of methods used for comparative | | | | | |

| <p>analyzes, their advantages, disadvantages and scope</p> <p>PEK_Knowledge of the contemporary methods of structure prediction</p> <p>Relating to skills:</p> <p>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</p> <p>PEU_U02 Ability to search sequence databases for similar sequences, including searches with similarity profiles, to identify homologs</p> <p>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</p> <p>PEU_U04 Ability to automate common bioinformatics tasks and analyses using scripting programming languages and specialized libraries</p> <p>Relating to social competences:</p> <p>PEU_K01 Ability to interpret, critically assess and communicate the results of bioinformatics analyzes</p> | | |
|--|---|-----------------|
| PROGRAM CONTENT | | |
| Lectures | | 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 similarity profiles. | 2 |
| Lec 7 | Introduction to Bayesian statistics and interpretation of information encoded in biological 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 individual tasks, with analysis of results | | |
| EVALUATION OF SUBJECT LEARNING OUTCOMES ACHIEVEMENT | | |
| Evaluation (F – forming (during semester), C – concluding (at semester end)) | Learning outcomes number | Way of evaluating learning outcomes achievement |
| F1_Lab | PEK_U01 | Individual task #1 |
| F2_Lab | PEK_U02 | Individual task #2 |
| F3_Lab | PEK_U03+PEK_U05 | Individual task #3 |
| F4_Lab | PEK_U04+PEK_U05 | Individual task #4 |
| C_Lab | | F1_Lab+F2_Lab+F3_Lab+F4_Lab (total score: 30 points) Score Grade 15-17,5 3,0 18-20,5 3,5 21-23,5 4,0 24-26,5 4,5 27-30 5,0 |
| C_Lec | PEK_W01, PEK_W02, PEK_W03, PEK_W04 | Exam |
| PRIMARY AND SECONDARY LITERATURE | | |
| <u>PRIMARY LITERATURE:</u> | | |
| [1] S.Q. Ye, Bioinformatics. A practical approach, Chapman & Hall/CRC, 2008 [2] I. Eidhammer, I. Johanssen, W.R. Taylor, Protein Bioinformatics - an algorithmic 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