

BIOINFORMATICS (BF-19),
UNDERGRADUATE STUDY PROGRAMME, FIRST BOLOGNA CYCLE
COURSE DESCRIPTIONS

COMPULSORY COURSES

COMPULSORY COURSES OF THE 1ST YEAR OF STUDY

Course name: **ALGEBRA II – LINEAR ALGEBRA**

Number of ECTS credits: **6**

Content:

Description of content:

- Groups, rings, domains. Polynomial rings.
- Vector space. Subspaces, linear operators. Linear independence. Basis and the dimension vector space.
- Eigenvalues. The characteristic and minimal polynomial.
- Inner product. Orthogonal systems. Gramm-Schmidt orthogonalization process. Norm. Norm of the matrix and the operator. Normal and related operators.
- Convexity in the vector space. Normed vector spaces as metric spaces. Isometries of \mathbb{R}^2 and \mathbb{R}^3 .

Course name: **ALGEBRA I – MATRIX CALCULUS**

Number of ECTS credits: **6**

Content:

- Vectors, analytic geometry in space.
- Matrices. Types of matrices and basic operations with matrices. Rank of a matrix. Inverse.
- Systems of linear equations. Matrix interpretation and the solvability theorem. Elementary matrices, Gauss method. Determinants. Cramer's rule.

Course name: **ANALYSIS II – INFINITESIMAL CALCULUS**

Number of ECTS credits: **6**

Content:

- Derivative. Mean value theorems. Differentiation of monotone functions. L'Hopital's rule. Higher derivatives. Taylor's formula. Local extrema. Convex and concave functions. Inflection points. The tangent method of finding the zeros.
- Indefinite Integral. Definite Integrals. Darboux and Riemann sums. Leibniz-Newton formula. Mean value theorems. Integration methods. Application of the definite integral in geometry. Improper integral. Numerical integration.
- The logarithm, the number e and the definition of exponentiation with the real exponent.
- Drawing planar curves.
- Sequences and function series. Power series. Taylor series. Elementary complex functions.

Course name: **ANALYSIS I – FOUNDATIONS OF ANALYSIS**

Number of ECTS credits: **6**

Content:

- Natural numbers. Rational numbers. Real numbers. Complex numbers.

- Sequences of real numbers. Limits and accumulation points. Cauchy condition. Upper and lower limit. Monotone sequences. Bolzano-Weierstrass theorem.
- Series. The convergence criteria. Absolutely and conditionally convergent series.
- Functions of real variables, even and odd functions, periodicity. Limits of functions, left and right limits. Continuity. Continuous functions on closed intervals limited. Bisection method for finding zeros.
- Elementary functions. Cyclometric functions.

Course name: **GENETICS**

Number of ECTS credits: **6**

Content:

- Fundamentals of transmission genetics: laws of inheritance, pedigree analysis of inheritance of dominant and recessive genes, molecular basics of dominant and recessive mutations.
- Autosomal and sex linked inheritance.
- Cytoplasmic inheritance.
- Basic laboratory methods in molecular genetics:
 - restriction enzymes, separation of DNA fragments, cloning DNA vectors, DNA recombinant techniques, methods for sequencing DNA, understanding and importance of genomic projects
- Analysis of genes, reverse genetics, methods for analysis of expressed sequences
- Morphology and structure elements of Eukaryote chromosome: telomere, centromere, role, replication model of telomere, organization of DNA on the chromosome, construction of chromatin, heterochromatin
- Evolution of genomes
- Organization of Eukaryote genome
- Genes and organization of genes
- Organel genomes, evolution and properties
- The pathway from DNA to protein
- Synthesis and processing of RNA
- Expression of genes
- Modification of genetic material; mutations on the chromosome, gene, genome. The causes of mutations, the impact on the organism and the importance in the evolution
- Repair mechanisms
- Recombinations
- Mobile genetic elements and their role
- Forensic genetics, applications
- Application of genomics

Course name: **ORGANIC CHEMISTRY AND BIOCHEMISTRY**

Number of ECTS credits: **6**

Content:

Organic chemistry:

- Importance and role of organic chemistry
- Structural characteristics of organic compounds
- Nucleophilic reactions
- Radical reaction
- Oxidations
- Reductions
- Carbohydrates
- Amino acids, peptides and proteins
- Isolation of organic compounds

Biochemistry:

- Principles of biochemistry
- Structure and function of biomolecules and catalysis: amino acids, proteins, enzymes, carbohydrates, DNA structure, lipids

- Bioenergetics and metabolism (catabolism and anabolism): glycolysis, krebs cycle, glyoxylate cycle, phosphogluconate pathway, oxidative phosphorylation, beta-oxidation, urea-cycle, gluconeogenesis, fatty acid biosynthesis, amino acid biosynthesis
- Integration, coordination and specificities of metabolism of organs

Course name: **DATA PROGRAMMING**

Number of ECTS credits: **6**

Content:

1. Introduction to data programming.

Programming languages for data programming, basic data structures and operations, applications.

2. Introduction to Python.

Interpreter environment VS code, syntax, numbers, strings, lists, files, sequential control, loop control, functions, lambda expressions, higher-order functions, functional programming.

3. Data structures in Python.

Lists, products, sequences, sets, dictionaries, loops for data structures.

4. Modules and classes in Python.

Syntax of modules, module environment, standard modules, syntax of classes, classes as objects, instances, methods, inheritance, iterators, generators.

5. Numeric Python (numpy).

N-dimensional arrays, indexes and array slicing, operations, data processing, data cleaning, data transformations, aggregations, grouping, timeseries, structured arrays.

6. Basics of R programming language.

History of R. Characteristics of R programming language, usage, graphical user interfaces for R, R studio - how it looks like and how it works, basic usage of command-line interface (practical example), numbers, vectors, R commands, functions and definitions, objects, attributes, matrices, lists, factors, data frames.

7. Reading and writing in R.

Reading of objects, vectors and matrices, reading data frames, reading from files, data generating, writing R objects, writing own simple functions in R.

8. Data managing in R.

databases, statistical models in R, graphical procedures, practical example of statistical model and graphic display usage.

9. Reproducible research with Rmarkdown, sharing reports on Rpubs and creating presentations.

10. Example of application in R with Shiny

Course name: **PROGRAMMING I**

Number of ECTS credits: **6**

Content:

- Basic building blocks of a computer program.

- Basic syntax of the programming language Java: Variables, types and expressions. Basic I/O operations. Decision statements. Control structures. Functions and parameters. Programs. Structural decomposition.

- Basic data structures: Simple types. Arrays. Records. Characters and strings. Data representation in computer memory. Memory allocation. Linked structures. Stack. Queue. List. Tree.

- Algorithms and problem solving: What is an algorithm? Problem solving strategies. The role of algorithms in problem solving. Algorithm implementation strategies. Debugging. Recursion – recursive functions, divide-and-conquer principle, backtracking, implementation of recursion.

- Programming languages overview: Types of programming languages. Flow control. Functions. Subprograms. Namespaces.

- Declarations and types: Types. Declarations of types. Safe typing. Type checking. Subtypes. Classes. Polymorphism.

- Abstraction mechanisms: Data abstractions. Simple types. Composite types. Flow abstractions. Subprograms and functions. Abstract data types. Objects and classes. Patterns. Modules.

Course name: **COMPUTER PRACTICUM**

Number of ECTS credits: **6**

Content:

The faculty network and basic usage rules:

- Description of the faculty computer network, login methods, password changing procedure, e-mail and mailing list usage, access to e-materials.

OS Linux basics:

- Description of the Linux OS and its many flavors, BASH shell usage basics.

Programming language C:

- The syntax of the C programming language and its usage to solve simple problems.

Course name: **THEORETICAL COMPUTER SCIENCE I**

Number of ECTS credits: **6**

Content:

- Fundamentals of mathematical theory, propositional calculus, truth tables, predicate calculus.
- Formal languages.
- Basic concepts of mathematical logic.
- Ways of describing sets. Basic relations between sets, basic operations on sets or families of sets. Power set. Relations. Graphs. Equivalence relations. Partial and linear ordering. Lattices and Boolean algebra. Well-ordered sets. Functions. Special types of functions. Categories.
- Finite and infinite, countable and uncountable sets.
- Cardinal and ordinal numbers. Peano arithmetic, mathematical induction.
- NBG and ZFC systems of axioms of set theory. Axiom of choice. Zorn's lemma.
- Fundamentals of symbolic computation (Mathematica).

COMPULSORY COURSES OF THE 2ND YEAR OF STUDY

Course name: **ALGORITHMS IN BIOINFORMATICS**

Number of ECTS credits: **3**

Content:

- Introduction to algorithms for biological sequence analysis: based on pattern searching, sliding windows, matrices, classical statistical methods and graphical presentations.
- Freshening of the concepts of automata theory, languages and computation.
- Searching biological sequences using finite-state automata and regular expressions
- Data structures for representing and searching biological sequences (suffix trees and arrays).
- Aligning sequence pairs - problem definition and methods: dot-matrix, dynamic programming (global and local alignment, alignment with gaps), heuristic approaches (FASTA, BLAST).
- Clustering: hierarchical and nonhierarchical methods.
- Aligning multiple sequences - heuristic approaches: generalized dynamic programming, 2-approximation algorithm, progressive methods (Feng-Doolittle, Clustal), Hidden Markov models (Viterbi).
- Reconstruction of phylogeny trees from distance matrices (clustering, additive phylogeny) and sequence alignment (small and large parsimony problem, Sankoff's algorithm).
- Sequence assembly: Hamiltonian and Eulerian path approaches.

Course name: **ANALYSIS III – FUNCTIONS OF MANY VARIABLES**

Number of ECTS credits: **6**

Content:

- Metric spaces. Cauchy-Schwarz inequality. Open and closed sets. Compactness and connectedness. Sequences in metric spaces. Cauchy sequences and complete spaces. Continuity

and uniform continuity. Properties of continuous mappings.

- Functions of several variables. Continuity, partial differentiability. Differential mapping from \mathbb{R}^n to \mathbb{R}^m . Jacobian matrix. Chain rule.
- Higher order partial derivatives. Taylor's formula. Theorem on the inverse and implicit functions. Local extrema and constrained local extrema
- Double and Multiple integrals. Properties. Theorems on existence. The substitution of new variables. Calculation and application.
- Proper and improper integrals with the parameter, Beta and Gamma functions. Stirling formula.

Course name: **NUCLEOTIDE SEQUENCE ANALYSIS**

Number of ECTS credits: **3**

Content:

- Applications of next generation sequencing (NGS)
- Review of NGS sequencing technologies, laboratory sample preparations for sequencing
- Databases with data from high-throughput technologies (SRA, Bioproject, TSA, WGS, ENA...)
- File formats (FASTQ, SAM/BAM, VCF)
- General approaches for genome sequencing (whole genome sequencing, exome sequencing, target sequencing)
- NGS data analysis workflow (quality assessment of NGS sequences, assembly, mapping, gene prediction and annotation)
- Tools for assembly of short reads de novo or mapping reads on a reference sequence (BWA, Bowtie2, TopHat,...)
- Tools for annotation (RepeatMasker, Splign, Genescan...)
- Visualization of NGS data (npr. Tablet, Genome workbench)
- Examples of other bioinformatic approaches, e.g. identifying tandem repeats, variants, development of molecular markers (SSR, SNP), microRNA analysis, genotyping by sequencing (GBS).

Course name: **INTRODUCTION TO DATABASE SYSTEMS**

Number of ECTS credits: **6**

Content:

- Introduction
- Entity-Relationship model
- Relational model
- Relational algebra and calculus
- SQL and QBE
- Disks and files
- Indexes
- Evaluation of relational operations
- Query optimization
- Transactions
- Concurrency control and crash recovery
- Database design

Course name: **INTRODUCTION TO MACHINE LEARNING AND DATA MINING**

Number of ECTS credits: **6**

Content:

Machine learning techniques:

- From basic to more advanced machine learning methods such as: decision trees, decision rules, regression, support vector machines, association rules ...
- The use of these techniques on real-life problems
- Building and evaluation of prediction and description models from the data.

Data mining:

- The use of machine learning techniques in the frame of the data mining process.
- Description of the CRISP methodology – data mining as a process consisting of the following steps: problem understanding, data understanding, data preparation, modeling, evaluation and deployment.
- The use of existing, open source data mining tools (e.g.: WEKA, Orange, A-Priori ...)

Course name: **DATA STRUCTURES AND ALGORITHMS**

Number of ECTS credits: **6**

Content:

The basic mathematical tool

- The evaluation function O , Ω , Θ and the differences between them
- What is the complexity of the problem and what is the complexity of solutions
- probability and randomness
- computational models: Pointer Machine, RAM, parallel computer

Basic data structures

- implicit data structures: array, stack, queue, heap
- explicit data structures: linked list, tree
- problems with concurrency

The basic abstract types and their performance

- Dictionary: hash tables, AVL, red-black, B – trees
- priority queue: heap, VEB
- ranking and selection

Sorting and friends:

- selection sort, merge sort, quicksort
- sorting in linear time
- sorting on parallel machines
- median and the k^{th} element

Basic algorithmic techniques

- greedy algorithms
- divide and conquer
- dynamic programming

Algorithms on graphs and networks

- search in width and in depth
- connected components
- minimum spanning trees
- shortest paths in graphs
- network flow
- parallel algorithms on networks with a view of the Internet

Selected algorithms

- mathematical algorithms with a view on cryptography: matrix multiplication, solving systems of equations, FFT, finding the greatest common divisor, modular arithmetic, exponents
- algorithms on strings with a view on bioinformatics: finding a substring
- for all problems we shall look at some basic parallel solutions

Course name: **PROGRAMMING II – CONCEPTS OF PROGRAMMING LANGUAGES**

Number of ECTS credits: **6**

Content:

- Introduction
- Lambda calculus
- Syntax
- Basic structures
- Functional languages
- Imperative languages

- Types
- Modules
- Objects and classes

Course name: **STATISTICS**

Number of ECTS credits: **6**

Content:

- Introduction. What is statistics? Examples of applications of statistical methods in natural sciences. Population. Sample.
- Random variables. Continuous and discrete random variables. Examples of random variables. Probability function, probability density function, cumulative distribution function.
- Descriptive statistics. Frequency distribution. Measures of central tendency. Quantiles. Measures of variability.
- Sampling. Introductory examples. Random sampling. Sampling distribution. Standard error. Confidence intervals.
- Graphical methods of data representation. Histogram. Scatter plots. Box and whisker plot. QQ-diagram.
- Hypothesis testing. Null hypothesis, alternative hypothesis. Errors in hypothesis testing. Examples: t-tests (mean of single sample, means of two independent/dependent samples, ANOVA)
- Correlation and dependence. Regression. Linear regression. Measures of linear dependence. Correlation coefficients. Non-linear dependencies. Association, contingency tables.

Course name: **FOUNDATIONS OF PHYSICS WITH BIOPHYSICS**

Number of ECTS credits: **6**

Content:

The course introduces basic concepts of physics which are essential for the understanding of the processes present in living organisms. The importance of these concepts is illustrated by their use in the explanation of specific cases.

Principal general themes of physical fundamentals:

- Intensive and extensive (additive) physical quantities and operation procedures in the measuring of these quantities (added to the basic ones will be the concept of density, pressure, concentration, temperature, potential, conductivity, optical properties, ...)
- Forms of energy (kinetic, potential, elastic, internal, chemical, electrical,...) and methods for increasing, decreasing and changing of energies (work, heat).
- Matter, electrical and heat flows and conservation laws.
- Oscillations and waves (mechanical and electromagnetic), and phenomena related to oscillations (reflection, refraction, interference).

More detailed themes of biophysical applications:

- Biomechanics (levers in the human body, biomechanics of human limbs, torsion deformations, body stability).
- Differences in pressure and concentration as a cause for matter flows (flows in plants, arteries and heart, exchange of gasses, osmosis processes).
- Temperature differences as a cause of heat flows (fundamentals of processes in living creatures with varying and fixed body temperatures, sources of energy and energy transformation).
- Differences of potential as a cause of electrical currents (a cell as the battery of electrochemical energy: the passing of matter through the cell membrane, action potential, electric field in a cell, passive and active electrical properties of the cell membranes).
- Vocal chords as a transmitter of sound and the biophysical image of the ear as a sound receptor.
- Optical instruments (camera, projector, microscope, and the eye as a light receptor).
- Basics of the molecular biophysics (properties of the basic components of biological systems and connections between them).

Course name: **INTRODUCTION TO BIOINFORMATICS**

Number of ECTS credits: **6**

Content:

BIOLOGICAL DATABASES

- Biological databases with data on a level of DNA, RNA, proteins, biological processes... (GenBank, EMBL-Bank, DDBJ, RefSeq, dbEST, Uniprot, KEGG, Reactome, itd.)
- Nucleotide / protein sequence file formats (GenBank, EMBL, GCG, FASTA,...), conversion between different file formats
- Access to information (Entrez, EDirect, SRS, dbfetch,...) and genome browsers (UCSC, Ensembl, NCBI GDV))

METHODS FOR NUCLEOTIDE/ AMINO ACID SEQUENCE ANALYSIS

- Alignment with different scoring matrices (PAM, BLOSUM)
- Local alignment (BLAST), global alignment
- Pairwise alignment with dotplots
- Methods for searching distantly related proteins (PSI-BLAST, DELTA-BLAST, HMMER)
- Multiple sequence alignment (MUSCLE, MAFFT, T-COFFEE, ClustalW, ProbCons)
- Tools for searching databases (e.g. NCBI BLAST, SIM), using different alignment methods and interpretation of the results

CHROMOSOME STRUCTURE ANALYSIS

- Tools for identification of regions with pseudogenes, repetitive DNA, transposons,...
- Tools for the prediction of genes, introns and exons, promoter regions, etc.

DNA SEQUENCING

- Sequencing with Sanger method
- Electropherogram analysis
- Base quality parameters (Phred value, KB basecaller Q value)
- Problems with sequence determination based on electropherograms (frameshift mutations, dye blobs, mixed base, etc.)

COMPULSORY COURSES OF THE 3RD YEAR OF STUDY

Course name: **BIOTECHNOLOGY**

Number of ECTS credits: **3**

Content:

- History of biotechnology
- Recombinant DNA technology
- Properties of expression vectors and cloning vectors
- RNA interference technology
- Genome editing methods (site-specific endonucleases CRISPR-Cas, TALEN, ZFN, meganucleases, oligonucleotide directed mutagenesis based on homology recombination process)
- Tools for developing of DNA construct based on the chosen mutagenesis method
- Methods for indirect / direct transfer of DNA / ribonucleoprotein complex in the plant or animal cell
- Methods for detection of construct integration into the genome or for the identification of the mutagenesis successfulness
- Procedures for development of transgenic plants
- Procedures for development of transgenic animals
- Gene therapy
- Practical examples of using biotechnology methods from the scientific literature
- Importance and achievements of biotechnology on various fields
- Legislation on the field of biotechnology

Course name: **EVOLUTIONARY BIOLOGY**

Number of ECTS credits: **6**

Content:

Historical development of evolutionary thought before and after Darwin-Wallace theory; Definition of life; Spontaneous variability (mutations - molecular variability and its significance for evolution); Natural selection; Evolution and biogeography; Evolution and speciation; Biological species concept; Evolution through time (with basics of palaeontology); Origin and evolution of genus *Homo*; Origin of life; Creationism and evolutionism.

Course name: **EVOLUTIONARY AND POPULATION GENETICS**

Number of ECTS credits: **6**

Content:

- The concept of molecular evolution
- Neutral theory
- Evolutionary changes in nucleotide sequences (models, genetic distance)
- Molecular clock
- Molecular phylogenetics
- Introduction to population genetics
- Genetic and phenotypic diversity (maintenance of genetic diversity, allele frequencies, genotype frequencies, heterozygosity)
- Molecular methods and sampling in population genetics
- Molecular markers: nuclear, chloroplast, mitochondrial
- Presentation of molecular markers and their use in population studies: RFLP, SNP, AFLP, RAPD, SSR, ISSR, DArT, RAMP, SRAP
- Evolutionary factors and their effects on genetic structure of populations: natural selection, mating, fragmentation of populations, migration (random mating, inbreeding, inbreeding depression, inbreeding coefficient, heterozygosity within populations, Wright's F statistic, models of migration)
- Parameters of genetic diversity
- Hardy-Weinberg equilibrium Linkage Disequilibrium analysis, causes of disequilibrium
- Random gene flow (genetic drift), Wright-Fisher model, inbreeding
- Irreversible and reversible mutation, influences on random gene flow, probability of fixation of new neutral mutations
- Genetic structure of populations (diversity, distribution of diversity between populations, genetic distances), population size, gene flow, reproduction, natural selection
- Molecular population genetics (sequence divergence, amino acid substitutions, nucleotide substitutions and their models, molecular clock), intraspecific polymorphism, polymorphism and divergence of non-coding regions, local recombination rates
- Phylogenetic trees
- The use of population genetics in species conservation
- The use of population genetics in agricultural science
- The use of statistical methods in population genetics

Course name: **PHYSICAL CHEMISTRY WITH CHEMINFORMATICS**

Number of ECTS credits: **6**

Content:

- Gasses and intermolecular forces : the ideal gas, kinetic theory of gases, intermolecular interactions, the van der Waals gas
- Energy: the first law of thermodynamics, changes of internal energy and enthalpy in physicochemical processes, thermochemistry
- Entropy: the second law of thermodynamics, entropy of gasses and gas mixtures, entropy changes in physicochemical processes, the third law of thermodynamics
- Biomolecular simulations: topology of biomacromolecules, protein data bank, force-field parameters, solvation and ionic atmosphere, energy minimization and molecular dynamics

- Structural analysis of trajectories: radial pair-distribution functions of ions, root-mean-square deviation of structure, root-mean-square fluctuation of structure, radius of gyration, heat-to-tail distance, hydrogen-bond analysis, conformational clustering
- Energy analysis with free energy: van der Waals and electrostatic interaction, energy decomposition
- Visualization of simulation trajectories and biomolecular computer graphics

Course name: **FUNCTIONAL GENOMICS**

Number of ECTS credits: **3**

Content:

In this course we will learn bioinformatics methods and approaches for analysis of functional genomics data and their application in the field of molecular biology and medicine.

Modern molecular biology techniques enable us to carry out measurements of a large number of genes and/or proteins simultaneously. We will learn approaches that are adapted for the analysis of high-throughput omics data. The emphasis will be on analysis of gene expression data using DNA-microarray and next generation sequencing (NGS) technologies.

We will learn statistical approaches and unsupervised and supervised machine learning methods that are adopted for analysis, integration, visualization and interpretation of omics data, as well as methods for discovery of relationships in omics data.

We will wrap up the course with a seminar, which will allow students to apply the presented approaches on transcriptome data from research studies of molecular mechanisms of various pathologies and diseases, such as metabolic syndrome, hepatocellular carcinoma, pulmonary fibrosis, etc. The seminar will be based on up-to-date research articles where the associated transcriptome data are publicly available in Gene Expression Omnibus (NCBI) or ArrayExpress (EMBL-EBI) database. The students will have an opportunity to extend the seminar for their final thesis.

Specific topics, methods and approaches which will be presented:

- Overview of high-throughput omics technologies for functional genomics research with emphasis on DNA-microarrays and NGS technologies.
- Characteristics of omics data at the field of functional genomics research with the emphasis on transcriptomic data.
- Design of experiments using DNA-microarray and NGS technologies.
- Pre-processing and quality control of DNA-microarray and NGS data.
- Statistical modeling of transcriptomic data (t-test, ANOVA, linear models, permutation methods, multiple comparisons adjustment) and variance-stabilization approaches.
- Supervised and unsupervised machine learning approaches and discovery of relationships in transcriptomic data.
- Assessing the quality of prediction models based on transcriptomic data (learning/testing sets, leave-one-out, cross validation).
- Resources and databases for annotation of genes and establishing sets of genes.
- Testing enrichment of gene sets.
- Visualization and interpretation of transcriptome data.

Course name: **MATHEMATICAL MODELLING IN BIOINFORMATICS**

Number of ECTS credits: **3**

Content:

- *Introduction.* What is mathematical modeling? The role of mathematical models in natural sciences. Types of mathematical models.
- *Programming tools.* A short overview of Octave/Scilab.
- *Differential equations and systems of differential equations as mathematical models in natural sciences.* Motivational examples. Equilibrium. (Linear) Stability of equilibria. Phase portraits. The basics of Poincare-Bendixon theory. The basics of bifurcation theory. Applications: epidemic models, models of competition, models of symbiosis, predator-prey dynamics, molecular kinetics, basic neurological models.

Course name: **SYSTEMS III – INFORMATION SYSTEMS**

Number of ECTS credits: **6**

Content:

Modern society is based on information management, which enables it to more quickly adapt to changes in needs and requirements. Computer aided information systems play the key role in this process. This course deals with their influence on the business environment as well as with their design and development.

Contents:

- Basic definitions and the role of information systems in information society.
- Information systems in business process.
- Systems development life cycle (SDLC)
- Systems analysis and design, structured and object-oriented approach.
- Data and process models and diagramming techniques. Data dictionary, structure diagram, DFD, ER, ELH. Introduction to UML.
- Collecting information, traditional and modern techniques.
- Interface design, service oriented architecture (SOA).
- Cloud computing.

Course name: **STRUCTURE OF BIOLOGICAL MOLECULES**

Number of ECTS credits: **3**

Content:

The first part of the course will be focused on an introduction to structures of biological molecules with subsequent presentation of more complex biological molecules. Then experimental methods important for structure identification will be introduced, especially protein crystallography and protein NMR. A review will be provided of main available databases and analytical tools for studying biological molecules (students will use these practically for analysing gene, transcripts, secondary, tertiary, quaternary structure, function, posttranslational modifications, interactions etc. for their chosen protein). In the course also theoretical methods for 3D structure prediction such as homology modelling fold recognition method and ab initio methods will be presented. Finally, students will briefly familiarize themselves with the utilization of computational chemistry for studying structure and modulator of biological modulators and with QSAR models for interaction of xenobiotics (environmental toxins, drug) and target biological molecules responsible for effects in organism (briefly properties of xenobiotics influencing their effects will be introduced).

Major themes:

- basic definition of biological molecules
- determination of protein structures: X-ray crystallography, NMR, electron microscopy, crystallization of proteins
- molecular modeling of protein structure
- QSAR relationship– structure/effect model of xenobiotics effects to biological macromolecules
- databases of biological molecules and their use (PDB, SCOP2, CATH, CSD, Expasy, GeneCards, UniProt, KEGG)
- computational tools for studying protein structure and selected examples

Course name: **SEMINAR (FINAL PROJECT PAPER)**

Number of ECTS credits: **6**

Content:

The course instructor, together with other teachers of the Bioinformatics study programme, presents potential thesis arguments, with relevant literature. Students are motivated to search for specific topics that they are interested in, and look for additional literature.

Presentation of the final thesis's structure and rules for in-text citations.

Final thesis provides a comprehensive overview of the chosen area belonging to the Bioinformatics study program. Student chooses a topic according to his interests and supervisor suggestions. The length and structure of the thesis is defined in the instructions for writing a thesis.

ELECTIVE COURSES

(Below you can find short descriptions of ALL elective courses of the study programme. In the table Elective courses you will find the list of the elective courses which were offered in the last two years.)

Course name: **BIOMOLECULAR MODELING**

Number of ECTS credits: **6**

Content:

Students will familiarize themselves with 3D macromolecular modelling from the available experimental data.

Major themes:

- macromolecular modeling by using crystal and NMR structures
- application of force fields in modeling and refinement of 3D structures
- dynamic characteristics of macromolecules, application and comparison of molecular dynamics and NMA (Normal Mode Analysis)
- visualization and interpretation of 3D models
- free-energy calculations
- molecular docking
- quantum-chemical simulations
- pharmacophore and virtual screening
- QSPAR

Computer exercises: practical examples of modeling of ligand-receptor structures using programs GROMACS, bio3d, VMD, MAIN, GAUSSIAN and ProBiS.

Course name: **MATHEMATICAL CHEMISTRY**

Number of ECTS credits: **6**

Content:

A. Theoretical basis

- Basics of classical combinatorics
- Basics of formal languages
- Basics of graph theory
- Matrices, eigenvalues, eigenvectors
- Maps on surfaces, Euler formula
- Graph algorithms in Sage and Python

B. Families of molecular graphs, Applications in chemistry and biology

- Molecules and chemical graphs
- DNA, RNA as strings
- Proteins as strings of aminoacids
- Trees, molecular trees, alkanes
- Benzenoids and their generalizations
- Fullerenes and fullerene-like structures
- Cubical and subcubical graphs
- Generating graphs using tools such as CaGe
- House of Graphs and similar databases

C. Problems on graphs with applications

- Spectral graph theory (Hückel theory, HOMO-LUMO)
- Matching theory (Kekulé structures)
- Hamilton cycles
- Graph coloring (Vertices correspond to atoms)
- Euler circuits (polyhedral self-assembly)
- D. Representation of graphs, graph drawing, molecular mechanics
 - Graph drawing using eigenvectors
 - Force-field methods
 - Rotagraphs and other polycriculants
- E. Graph invariants and topological indices
 - Wiener index
 - Szeged index
 - HL index
 - etc.

Each chapter is supported by programs in Sage.

Course name: **SYSTEMS BIOLOGY APPROACH IN HUMAN DISEASE STUDY**

Number of ECTS credits: **6**

Content:

In this course students will familiarize with critical parameters in modern experimental medical research.

Students will learn about developing good research questions and hypothesis, clinical studies and experimental design, variables and confounders, statistical analysis and data interpretation and presentation. The students will practice the selection of variables, consider the influence of co-variables, testing multiple hypothesis and the effect of sample size and power analysis. They will be introduced to basics of epidemiology and why it is important for human disease research, evidence based medicine and meta-analyses.

The seminar lectures will include the basics of molecular biology methods with an emphasis on presentation of data about genes, epigenetics, transcripts, proteins and metabolites and their functions and interactions. Students will practically use programs for analysis and graphical data presentation of data and learn on examples how to write a scientific paper or research thesis and how to make an oral research presentation. Students will learn how to answer concrete research questions in human disease. Seminar lectures will present methods on real-world data from studies of molecular mechanisms and diseases, such as immune response and metabolic syndrome.

Major themes:

- Formulating experimental question and testing hypothesis (variables- dependent, independent, co-variables, confounding; drawing conclusions-deduction, induction, comparison; validity, reliability of results-errors (type I, II, confidence interval, bootstrapping), population effect size and power analysis
- Study design, *in vitro/in vivo* experiments, types of pre-clinical, clinical studies; sampling, sample size, replication, randomization; error and bias of sample, experiment performance and study design
- Conducting multiple comparisons (post hoc tests, Bonferroni, Tukey..), ANCOVA, practical use of regression, basics of epidemiology
- Biostatistic of diagnostic test and method comparison (analytical, diagnostic specificity, sensitivity, predictive value)
- Systems medicine, 4P medicine, precision medicine
- Systems medicine in research and development of new drugs
- Ethical issues, informed consent, personal information privacy
- Overview of –omic data methods and methods for their further confirmation in medical studies with emphasis on transcriptomics, proteomics and metabolomics; overview of databases and freely available tools for their analysis with cases from different diseases

Practical approaches in solving modern medical research questions:

- Each seminar lecture will focus on one of human diseases and the data gathered about it (eg rheumatoid arthritis, atherosclerosis etc)
- Experimental data obtained from genomics, epigenetics, proteomics, metabolomics will be used and their advantages/disadvantages discussed
- Using online deposited data and data bases
- Data visualization, tables and graphs for research paper/thesis and oral presentation -using program for data presentation and statistical analysis
- Students will practice study conduct to compare two study populations and think of what variables could affect the treatment outcome or disease outcome and propose possible way of obtaining those information's from study population (i.e. digital biomarkers, increased use of wearable devices provide unique opportunities to develop innovative therapeutic strategies)

As INTERNALLY ELECTIVE COURSES students can also choose the following courses from the undergraduate study programmes at UP FAMNIT:

- **Mathematics:** [Introduction to Numerical Calculations; Algebra III – Abstract Algebra; Game Theory; Probability; Stochastic Processes I;](#)
- **Mathematics in Economics and Finance:** [Stochastic Processes II;](#)
- **Computer Science:** [Theoretical Computer Science II – Formal Languages and Computability; Theoretical Computer Science III – Information Theory; Systems I – Hardware; Systems II – Operating Systems and Computer Networks; Programming III – Concurrent Programming; Computer Networks;](#)
- **Conservation Biology:** [Conservation Biology; Systematic Botany and Geobotany; Environmental Monitoring; Geographical Information Science and Systems; Marine Biodiversity; Protected Areas and Sustainable Use; Biogeography; Ecotoxicology; Biodiversity and Ecology of the Mediterranean; Biological Topics in English; Ecology.](#)
- **Mediterranean Agriculture:** [Plant Biotechnology; Plant Molecular Diagnostic; Basics of Plant Production in the Mediterranean Area.](#)

Contents of the above mentioned courses are available in the presentation of the study programme (Course structure).