KAZAKH-BRITISH TECHNICAL UNIVERSITY

JSC «Kazakh British Technical University»
School of IT and Engineering

APPROVED BY
Dean of SITE
Azamat Imanbayev
« 06 » January 2025

SYLLABUS

Discipline: Introduction to Machine Learning
Term: Spring 2025
Instructors full name: Olivier Jaylet

Personal Information about the Instructor	Time and place of classes		Contact information	
	Lessons	Office Hours	Tel.	e-mail
Olivier Jaylet	According to the schedule	Tuesday, 16h – 17h	+33785991302	o.jaylet@kbtu.kz

COURSE DURATION: 5 ECTS, 15 weeks, 45 class hours

MS Teams code

GENERAL COURSE AIMS:

(1) To develop solid understanding of mathematical concepts which underpin the modern field of machine learning and artificial intelligence
(2) To acquire proficiency in Python programming using essential data science libraries
(3) To explore the main principles of contemporary machine learning

COURSE DESCRIPTION

This course provides an introduction to the fundamental concepts and techniques in the field of machine learning. Students will gain a solid foundation in the mathematical and statistical principles that underlie machine learning algorithms, as well as practical skills in Python programming using libraries such as NumPy, Pandas, Matplotlib and Scikit-learn.

COURSE OBJECTIVES

The objective of this course is to provide students with the fundamental knowledge and skills which will enhance their competence in the field of modern data science and machine learning.

COURSE OUTCOMES

At the end of the course and having completed the essential reading and activities students should be able to:

- use mathematical language to describe different problems from machine learning domain
- Solve mathematical problems that form the foundation of ML algorithm.
- Select and implement basic machine learning models and evaluate their performance.
- be well-prepared to delve deeper into more advanced machine learning concepts, algorithms, and applications

COURSE PREREQUISITES:

- Basic programming experience with Python
- Basic concepts of linear algebra: vectors and matrices and operations on them
- Single and multivariate calculus topics such as derivatives and integrals
- Key notions of probability and statistics

COURSE POST REQUISITES:

Knowledge and skills obtained during study of this course are used in following courses: Machine Learning, Deep Learning, Reinforcement Learning.

LITERATURE

1. Murphy K.P. Machine learning: A probabilistic perspective. MIT Press, second edition (2012)
2. Bishop C.M. Pattern recognition and machine learning. Springer (2006)
3. Hastie T., Tibshirani R. and Friedman J. H. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer (2009)
4. James D., Witten D., Hastie T., Tibshirani R. Taylor J. An Introduction to Statistical Learning with Applications in Python (2023)

COURSE ASSESSMENT CRITERIA

Assessment occurs continuously throughout the course. The evaluation will be based on the levels of (maximums in %):

Type of activity	Final scores
Attendance /participation	**10%**
Practice	**0%**
SIS	**20%**
Midterm exam	**15%**
Endterm exam	**15%**
Final exam	**40%**
Total	**100%**

TASKS

for students' independent study (SIS)

COURSE CALENDAR

Week	Class work			SIS
	Topic	**Lectures**	**Practice**	**SIS defense**
1	**Introduction to machine learning concepts.** Overview of machine learning and its applications. Types of machine learning. Python libraries for data analysis and machine learning.	1	2	
2	**Data and datasets.** Tabular, image, text data. Embeddings. Sources of datasets. Data mining with Python	1	2	
3	**Supervised learning.** Model, loss function, train and test error. Overfitting and underfitting. Classification and regression tasks.	1	2	
4	**Vectors and matrices**. Vector norms. Inner and outer product. Matrix product. Inverse Matrix. Rank of a matrix	1	2	SIS 1
5	**Linear regression.** Theoretical foundations. Simple Vs. Multiple linear regression. Polynomial regression. RSS and R2 scores	1	2	
6	**Linear algebra recap.** Rank of a matrix. Orthogonal projections. Least squares approximation, application to linear regression	1	2	SIS2
7	**Optimization.** Derivatives, gradients and differentials. Gradient descent. Newton's method. Integrals.	1	2	
8	Midterm Exam	1	2	
9	**Logistic regression.** Theoretical foundations. Representation of the model. Cost function and gradient descent. Multiclass classification. Regularization	1	2	
10	**Probability/Stats recap.** Distributions, densities, random variables. Bias-variance decomposition. Law of big numbers and central limit theorem.	1	2	
11	**Introduction to SVM & KNN.** Theoretical foundations. Maximal margin classifier. Optimization problem. Perfect separable margin. Support vector classifiers.	1	2	
12	**Introduction to Tree-based methods.** Decision **Trees. Regression trees. Binary splitting. Recursive algorithm. Classification Trees. Gini index & cross entropy**	1	2	SIS3
13	**Unsupervised learning / Clustering.** K-means. Distance metrics. Cluster separability. Initialization and local optima.	1	2	
14	**Evaluation Metrics.** Classification metrics: accuracy, precision, recall, AUC, confusion matrix, log loss. Regression metrics: MSE, MAE, R^2-score	1	2	

15	Endterm exam	1	2	
16-17	Final Exam			

Course assessment schedule

No	**Assessment criteria**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	
1	Attendance / participation	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	**10%**
2	SIS			*			*						*				**20%**
3	Midterm/ End-term								*							*	**30%**
4	Final exam																**40%**
	Total																**100%**

Class sessions – will be a mixture of information, discussion and practical application of skills.
Practice – will reinforst the students knowledge by practical appliance of lectured materials.
In-class assessment – will prepare students for their mid-term and final assessment and identify the competence level they have achieved on a related subject matter, the aim being to diagnose potential discrepancies in students' understanding and performance in order to make specific adjustments to the course content and procedures and/or to assign additional assignments to certain individuals or the whole group.
Home assignments – will consolidate the concepts and materials taken during in-class activities, help students to expand the content through diverse background resources and/or practise certain skill areas; they will also develop the students' ability to work individually in exploring and examining related issues.
SIS (Student Independent Study) – comprises group Project to be done by students on the independent basis. Students are supposed to use knowledge and skills acquired in class to do the project. Assistance and advice will be provided by teachers during office hours.
TSIS (Teacher Supervised Student Independent Study) – student self-made project.
End-term test – a diagnostic test used to identify the students' progress, their strengths and weaknesses, intended to force student to prepare for Final Exam. It includes computer-based test.
Final examination – 1) an attainment test designed to identify how successful the students have been achieving objectives.
Grading policy:
<u>Intermediate attestations</u> (on 8th and 15th week) join topics of all lectures, practice, laboratories, SIS, TSIS and materials for reading discussed to the time of attestation. Maximum number of points within attendance, activity, SIS, TSIS and laboratories for each attestation is 30 points.
<u>Final exam</u> joins and generalizes all course materials, is conducted in the complex form with quiz and problem. Final exam duration is 100 min. Maximum number of points is 40. At the end of the semester you receive overall total grade (summarized index of your work during semester) according to conventional KBTU grade scale.

<u>ATTENTION!</u>

1) If student missed more than **30% of lessons student receives «F (Fail)»** grade;

2) If for two attestations student receives 29 or less points, this student is not accepted to final exam and for all course he (she) receives **«F (Fail)» grade;**

3) If student receives on final exam 9.4 or less points, then independently on how many points he (she) received for two attestations, in whole he (she) receives **«F (Fail)» grade;**

In the case of missing or being late for final exam without plausible reason, independently on how many points he (she) received for two attestations, in whole he (she) receives **«F (Fail)» grade.**

ACADEMIC POLICY:

1. <u>Cheating, duplication, falsification of data, plagiarism are not permitted under any circumstances!</u>
2. Students must participate fully in every class. While attendance is crucial, merely being in class does not constitute "participation". Participation means reading the assigned materials, coming to class prepared to ask questions and engage in discussion.
3. Students are expected to take an active role in learning (the instructor will provide the information and guidelines to do this).
4. Students must come to class on time.
5. Students are to take responsibility for making up any work missed.
6. Make up tests in case of absence will not normally be allowed.
7. Mobile phones must always be switched off in class.
8. Students should always show tolerance, consideration and mutual support towards other students.

<table>
<tr><th colspan="2">Grade</th><th>Achievement percentage</th><th>Assessment criterion</th></tr>
<tr><td rowspan="2">«Excellent»</td><td>A</td><td>95-100%</td><td rowspan="2">This grade is given when the student:

demonstrated a complete understanding of the course material;
did not make any errors or inaccuracies;
completed control and laboratory work in a timely and correct manner, and submitted reports on them;
demonstrated original thinking;
submitted control quizzes on time and without any errors;
completed homework assignments;
engaged in research work;
independently used additional scientific literature in studying the discipline;
was able to independently systematize the course material.</td></tr>
<tr><td>A -</td><td>90-94%</td></tr>
<tr><td rowspan="3">«Good»</td><td>B+</td><td>85-89%</td><td rowspan="3">This grade is given when the student:

Has mastered the course material at no less than 75%;
Did not make gross errors in responses;
Timely completed control and laboratory work and submitted them without fundamental remarks;
Correctly completed and timely submitted control tests and homework assignments without fundamental remarks;</td></tr>
<tr><td>B</td><td>80-84%</td></tr>
<tr><td>B-</td><td>75-79%</td></tr>
</table>

	C+	70-74%	Utilized additional literature as indicated by the instructor; Engaged in research work, made non-fundamental errors, and fundamental errors corrected by the student themselves; Managed to systematize the course material with the help of the instructor.
«Satisfactory»	C	65-69%	This grade is given when the student: Has mastered the course material no less than 50%; Required assistance from the instructor when completing control and laboratory work, homework assignments; Made inaccuracies and non-fundamental errors when submitting control tests; Did not demonstrate activity in research work, relied solely on the educational literature indicated by the instructor; Experienced more difficulty in systematizing the material.
	C-	60-64%	
	D+	55-59%	
	D	50-54%	

Senier Lecturer **Olivier Jaylet**

Minutes # 9 of School of Information Technology and Engineering meeting on January 6, 2025