



Partner



SHAKHMARDAN
YESSENOV
FOUNDATION

YESSENOV DATA LAB SUMMER SCHOOL

June 11 – August 3, 2018
Almaty



2018



Table of Contents

About Yessenov Data Lab

Program stages

Who can apply for the program?

Apply to the Program

School's Syllabus

Week 1. Python

Week 2. Linear Models for Classification and Regression

Week 3. Working with Features (PCA, Classification)

Week 4. Neural Networks

Week 5. Deep Learning in Computer Vision and Reinforcement Learning

Solving Kaggle cases?

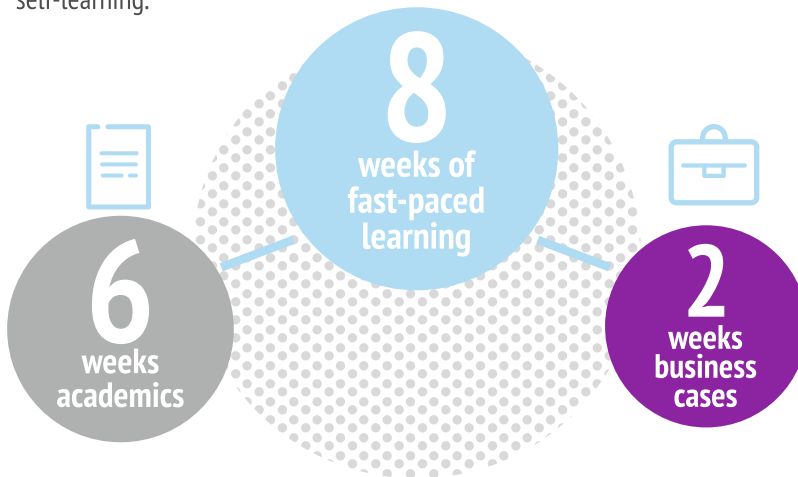
Week 6. Natural Language Processing (NLP)

Weeks 7-8. Project challenge



► About Yessenov Data Lab

The Yessenov Data Lab is an 8-week long intensive summer school that fast launches into the Data Scientist specialization. Participants solve the challenges businesses face and are equipped with knowledge to continue growing by self-learning.



School's dates: June 11 – August 3, 2018

Schedule: Mon-Fri, 9:00 am-6:00pm

Participants: 20 people

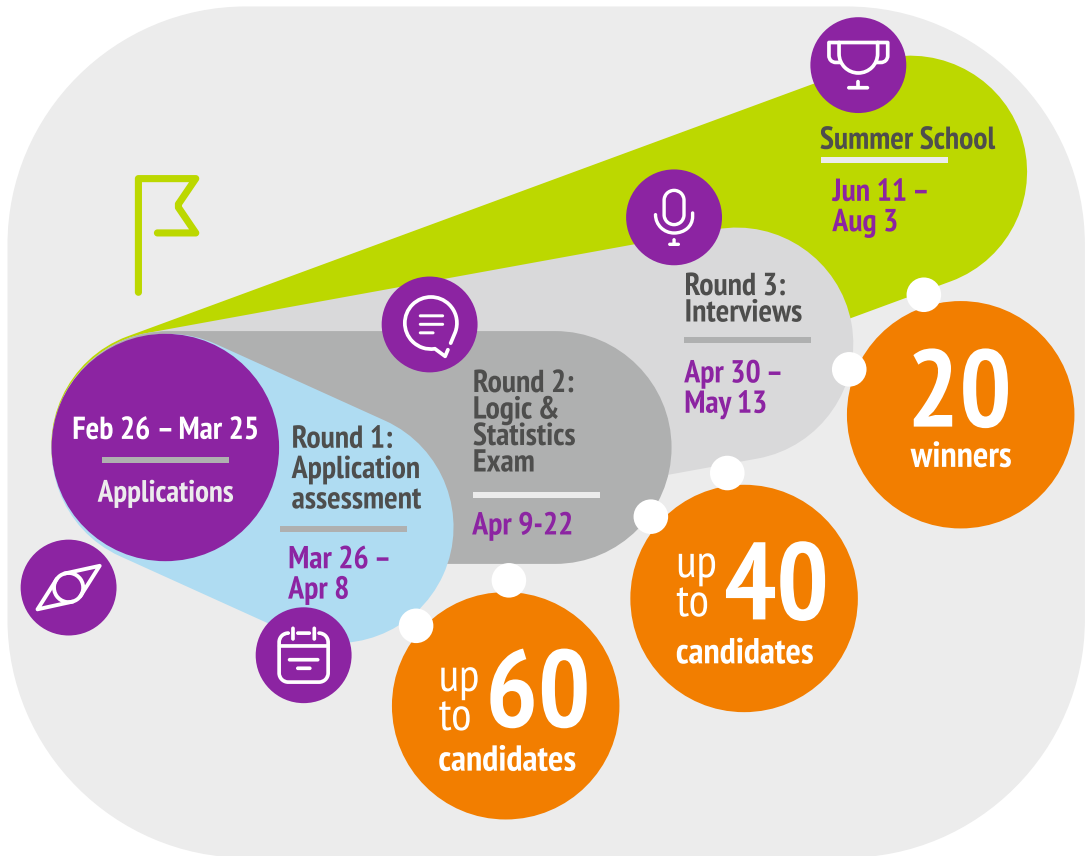
Venue:

Almaty Management University

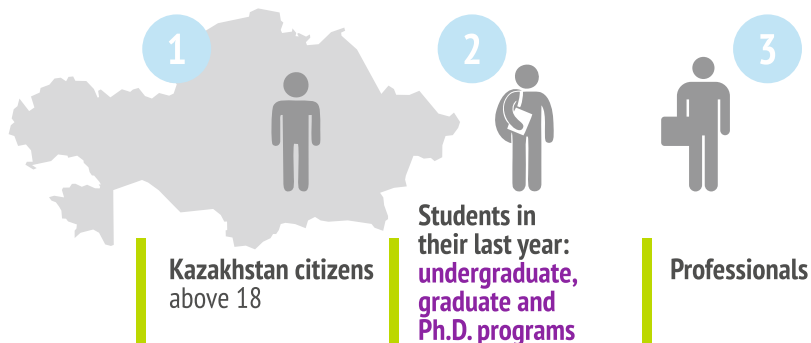
THE GRADUATES OF THE SUMMER SCHOOL CAN LOOK FOR TO ACQUIRE THE FOLLOWING SKILLS:

1. Programming in Python within data analysis
2. Preprocessing
3. Visualization of data and finding data dependencies
4. Forecasting based on historical data
5. Understanding different algorithms of training
6. Right choice of training model
7. Fundamental understanding of Neural Networks

► Program stages



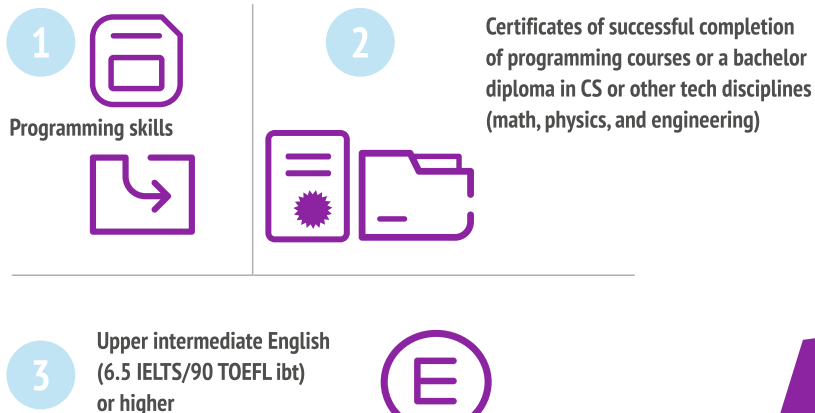
Who can apply for the program?



REQUIREMENTS FOR CANDIDATES:

- Strong analytical skills
- Basic knowledge of statistics and linear algebra
- Determination and result-oriented

THE FOLLOWING ARE A PLUS:



► Apply to the Program



ADDITIONAL DOCUMENTS LIST:

- 1. Application form
- 2. Copy of ID
- 3. Copy of diplomas, certificates on completion of courses (programming, statistics, etc.), participation in Olympiads (math, IT or any other tech disciplines)
- 4. Copy of transcript (all completed semesters) and a copy of bachelor degree diploma with transcript (for graduates and specialists)
- 5. Essay on “I want to learn data analysis to...”
- 6. Detailed portfolio demonstrating achievements in IT field (where possible)
- 7. Certificates of English language tests (where possible)



School's Syllabus

Week 1. Python

June 11-15

Day 1

- 09:00 – 10:00 Registration
- 10:00 – 11:30 What is Data Mining, Big Data? Examples
- 11:30 – 11:45 Coffee break
- 12:00 – 13:15 Case study: Titanic on Kaggle
- 13:15 – 14:30 Lunch
- 14:30 – 16:00 Python: Introduction. Variables, list, conditions, loops
- 16:00 – 16:15 Coffee break
- 16:15 – 18:00 Lab work: Basics of Python

Day 2

- 10:00 – 11:30 Data structures: list, sets, dictionaries
- 11:30 – 11:45 Coffee break
- 12:00 – 13:15 NumPy library: Introduction
- 13:15 – 14:30 Lunch
- 14:30 – 16:00 Lab work: data structures and NumPy
- 16:00 – 18:00 Team building

Day 3

- 10:00 – 11:30 Pandas and SciPy libraries: Introduction. Data upload
- 11:30 – 11:45 Coffee break
- 12:00 – 13:15 Grouping of data. Filters, sorting
- 13:15 – 14:30 Lunch
- 14:30 – 16:00 Lab work: CSV, TXT, QuandL
- 16:00 – 16:15 Coffee break
- 16:15 – 18:00 Lab work: CSV, TXT, QuandL

Day 4

- 10:00 – 11:30 Object-oriented programming
- 11:30 – 11:45 Coffee break
- 12:00 – 13:15 Case study: Coders Strike Back on codinggame.com
- 13:15 – 14:30 Lunch
- 14:30 – 16:00 Lab work: codinggame.com: simple tasks
- 16:00 – 16:15 Coffee break
- 16:15 – 18:00 Lab work: codinggame.com: Coders Strike Back

Day 5

- 10:00 – 11:30 Data upload. Data pre-processing
- 11:30 – 11:45 Coffee break
- 12:00 – 13:15 Simple visualization (2D Arrays)
- 13:15 – 14:30 Lunch
- 14:30 – 16:00 Lab work: Pandas
- 16:00 – 16:15 Coffee break
- 16:15 – 18:00 Lab work: Matplotlib



Kuanysh Abeshev
AlmaU



Timur Bakibayev
Professor AlmaU



School's Syllabus

Week 2. Linear Models for Classification and Regression

June 18-22

Day 1

10:00 – 11:30	Optimization, gradient decent method
11:30 – 11:45	Coffee break
12:00 – 13:15	Lab work
13:15 – 14:30	Lunch
14:30 – 16:00	Lab work:
16:00 – 16:15	Coffee break
16:15 – 18:00	Lab work

Day 2

10:00 – 11:30	Linear models for classification and regression
11:30 – 11:45	Coffee break
12:00 – 13:15	Lab work
13:15 – 14:30	Lunch
14:30 – 16:00	Lab work
16:00 – 16:15	Coffee break
16:15 – 18:00	Lab work

Day 3

10:00 – 11:30	Overfitting, generalization
11:30 – 11:45	Coffee break
12:00 – 13:15	Lab work
13:15 – 14:30	Lunch
14:30 – 16:00	Lab work
16:00 – 18:00	Team building

Day 4

10:00 – 11:30	Cross-validation
11:30 – 11:45	Coffee break
12:00 – 13:15	Lab work
13:15 – 14:30	Lunch
14:30 – 16:00	Lab work
16:00 – 16:15	Coffee break
16:15 – 18:00	Lab work

Day 5

10:00 – 11:30	Quality metrics
11:30 – 11:45	Coffee break
12:00 – 13:15	Lab work
13:15 – 14:30	Lunch
14:30 – 16:00	Lab work
16:00 – 16:15	Coffee break
16:15 – 18:00	Lab work



Dmitriy Rusanov
Data Scientist,
EPAM Systems



School's Syllabus

Week 3. Working with Features (PCA, Classification)

June 25-29



Michael Lipkovich
Lead big data engineer,
EPAM Systems

Day 1

10:00 – 11:30	Classification, decision tree and k-Nearest Neighbours
11:30 – 11:45	Coffee break
12:00 – 13:15	Lab work
13:15 – 14:30	Lunch
14:30 – 16:00	Lab work
16:00 – 16:15	Coffee break
16:15 – 18:00	Lab work

Day 2

10:00 – 11:30	Decision tree ensembles: bagging, boosting, random forest
11:30 – 11:45	Coffee break
12:00 – 13:15	Lab work
13:15 – 14:30	Lunch
14:30 – 16:00	Lab work
16:00 – 16:15	Coffee break
16:15 – 18:00	Lab work

Day 3

10:00 – 11:30	Unsupervised learning: PCA, clustering
11:30 – 11:45	Coffee break
12:00 – 13:15	Lab work
13:15 – 14:30	Lunch
14:30 – 16:00	Lab work
16:00 – 18:00	Team building

Day 4

10:00 – 11:30	Feature selection
11:30 – 11:45	Coffee break
12:00 – 13:15	Lab work
13:15 – 14:30	Lunch
14:30 – 16:00	Lab work
16:00 – 16:15	Coffee break
16:15 – 18:00	Lab work

Day 5

10:00 – 11:30	Support vector machine (SVM)
11:30 – 11:45	Coffee break
12:00 – 13:15	Lab work
13:15 – 14:30	Lunch
14:30 – 16:00	Lab work
16:00 – 16:15	Coffee break
16:15 – 18:00	Lab work



School's Syllabus

Week 4. Neural Networks

July 2-6

Day 1

10:00 – 11:30	Neural networks: Introduction. Perceptron
11:30 – 11:45	Coffee break
12:00 – 13:15	Back-propagation
13:15 – 14:30	Lunch
14:30 – 16:00	Lab work: Neural Network implementation
16:00 – 16:15	Coffee break
16:15 – 18:00	Lab work: Neural Network implementation

Day 2

10:00 – 11:30	Keras library: Introduction
11:30 – 11:45	Coffee break
12:00 – 13:15	Keras library: Introduction. Continued
13:15 – 14:30	Lunch
14:30 – 16:00	Lab work
16:00 – 16:15	Coffee break
16:15 – 18:00	Lab work

Day 3

10:00 – 11:30	Convolutional neural networks (CNN)
11:30 – 11:45	Coffee break
12:00 – 13:15	Lab work: image analysis
13:15 – 14:30	Lunch
14:30 – 16:00	Lab work: image analysis
16:00 – 18:00	Team building

Day 4

10:00 – 11:30	Recurrent neural network (RNN)
11:30 – 11:45	Coffee break
12:00 – 13:15	Lab work: text analysis
13:15 – 14:30	Lunch
14:30 – 16:00	Lab work: text analysis
16:00 – 16:15	Coffee break
16:15 – 18:00	Lab work: text analysis

Day 5

10:00 – 11:30	Problems of overfitting. Data augmentation
11:30 – 11:45	Coffee break
12:00 – 13:15	Lab work
13:15 – 14:30	Lunch
14:30 – 16:00	Lab work
16:00 – 16:15	Coffee break
16:15 – 18:00	Lab work



Marina Gorlova
Analyst,
Yandex Money



School's Syllabus

Week 5.

Deep Learning in Computer Vision and Reinforcement Learning. Solving Kaggle cases?

July 9-13

Day 1

10:00 – 11:30 MNIST, Fashion MNIST, LFW datasets classification
11:30 – 11:45 Coffee break
12:00 – 13:15 Lab work: work on an example
13:15 – 14:30 Lunch
14:30 – 16:00 Lab work: work on an example
16:00 – 16:15 Coffee break
16:15 – 18:00 Lab work: work on an example

Day 2

10:00 – 11:30 VGG, ResNet and Inception architectures. What neural networks see
11:30 – 11:45 Coffee break
12:00 – 13:15 Lab work: work on an example
13:15 – 14:30 Lunch
14:30 – 16:00 Lab work: work on an example
16:00 – 16:15 Coffee break
16:15 – 18:00 Lab work: work on an example

Day 3

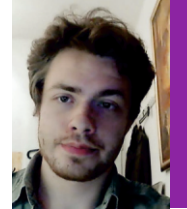
10:00 – 11:30 From classification to segmentation. Kaggle Challenges review
11:30 – 11:45 Coffee break
12:00 – 13:15 Lab work: work on an example
13:15 – 14:30 Lunch
14:30 – 16:00 Lab work: work on an example
16:00 – 18:00 Team building

Day 4

10:00 – 11:30 Autoencoders and Variational Autoencoders. Pose estimation
11:30 – 11:45 Coffee break
12:00 – 13:15 Lab work: work on an example
13:15 – 14:30 Lunch
14:30 – 16:00 Lab work: work on an example
16:00 – 16:15 Coffee break
16:15 – 18:00 Lab work: work on an example

Day 5

10:00 – 11:30 Reinforcement learning. Supervised learning limits
11:30 – 11:45 Coffee break
12:00 – 13:15 Lab work: work on an example
13:15 – 14:30 Lunch
14:30 – 16:00 Lab work: work on an example
16:00 – 16:15 Coffee break
16:15 – 18:00 Lab work: work on an example



Dmitriy Kotoenko
AGT International,
Computer Vision
Research Assistant



School's Syllabus

Week 6. Natural Language Processing (NLP)

July 16-20

Day 1

10:00 – 11:30	What is Natural Language Processing?
11:30 – 11:45	Coffee break
12:00 – 13:15	Lab work: NLTK library
13:15 – 14:30	Lunch
14:30 – 16:00	Lab work: NLTK library
16:00 – 16:15	Coffee break
16:15 – 18:00	Lab work: NLTK library

Day 2

10:00 – 11:30	Working with text corpus
11:30 – 11:45	Coffee break
12:00 – 13:15	Lab work: working with realistic example
13:15 – 14:30	Lunch
14:30 – 16:00	Lab work: working with realistic example
16:00 – 16:15	Coffee break
16:15 – 18:00	Lab work: working with realistic example

Day 3

10:00 – 11:30	Text classification
11:30 – 11:45	Coffee break
12:00 – 13:15	Lab work: working with realistic example
13:15 – 14:30	Lunch
14:30 – 16:00	Lab work: working with realistic example
16:00 – 18:00	Team building

Day 4

10:00 – 11:30	Information extraction from text
11:30 – 11:45	Coffee break
12:00 – 13:15	Lab work: work on an example
13:15 – 14:30	Lunch
14:30 – 16:00	Lab work: work on an example
16:00 – 16:15	Coffee break
16:15 – 18:00	Lab work: work on an example

Day 5

10:00 – 11:30	Sentiment analysis
11:30 – 11:45	Coffee break
12:00 – 13:15	Lab work: work on an example
13:15 – 14:30	Lunch
14:30 – 16:00	Lab work: work on an example
16:00 – 16:15	Coffee break
16:15 – 18:00	Lab work: work on an example



Daniyar Mukhtarhanuly
Alem Research

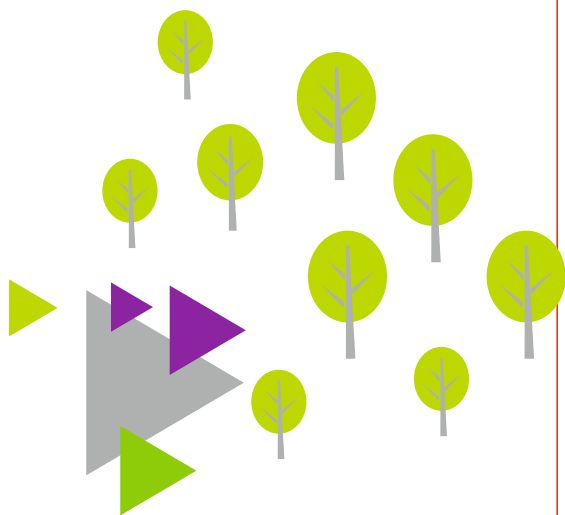


School's Syllabus

Week 7 и 8. Project challenge

July 23 – August 3

Kazakhstani companies that use data analysis will provide the program participants with challenges of real businesses. Successful graduates of the School will receive job offers.



STAY: yessenovfoundation.org
IN: info@yessenovfoundation.org
TOUCH: +7 727 346 92 88

