

ЛЕТНЯЯ ШКОЛА



**YESENOV
DATA LAB**

11 июня – 3 августа 2018
Алматы

В партнерстве с



Kaspi.kz



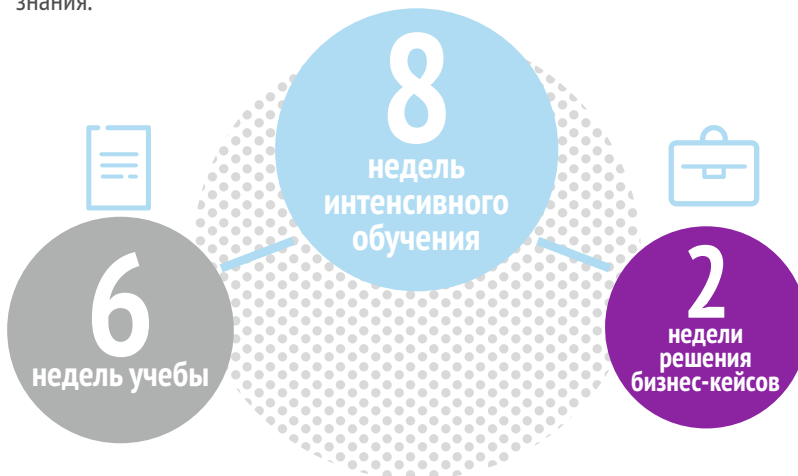
Оглавление

- Что такое летняя школа Yessenov data lab?
- Этапы программы
- Кто может участвовать в конкурсе?
- Участие в конкурсе
- Программа обучения
 - ▶ Неделя 1. Язык программирования Python
 - ▶ Неделя 2. Линейные модели классификации и регрессии
 - ▶ Неделя 3. Работа с признаками (PCA, классификация)
 - ▶ Неделя 4. Нейронные сети
 - ▶ Неделя 5. Глубокое обучение в компьютерном зрении и обучении с подкреплением. Решение Kaggle задач
 - ▶ Неделя 6. Обработка естественных языков (NLP)
 - ▶ Неделя 7 и 8. Проектные задачи



► Что такое летняя школа Yessenov data lab?

Летняя школа Yessenov data lab это 8-недельный интенсив, который позволяет за 8 недель ознакомиться с профессией аналитика данных, решать реальные задачи бизнеса и в дальнейшем самостоятельно совершенствовать полученные знания.



Даты проведения: 11 июня – 3 августа 2018

Расписание: по будням с 9:00 до 18:00

Количество участников: 20

Место проведения:

Almaty Management University

ВЫПУСКНИКИ ШКОЛЫ ПРИОБРЕТУТ НАВЫКИ:

1. программирования на языке Python в рамках анализа данных
2. предварительной обработки данных, подготовки их для последующего анализа (preprocessing)
3. визуализирования данных и нахождения в них зависимости
4. составления прогноза на основании исторических данных
5. понимания различных алгоритмов обучения
6. правильного выбора модели обучения
7. фундаментальное понимание работы нейронных сетей



► Этапы программы





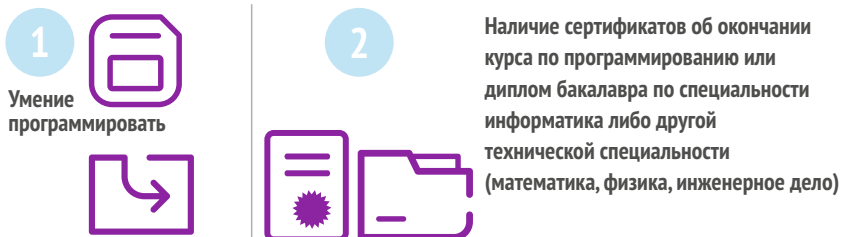
▶ Кто может участвовать в конкурсе?



ТРЕБОВАНИЯ К КАНДИДАТАМ:

- сильные аналитические способности
- знание основ статистики и линейной алгебры
- направленность на результативность и целеустремленность

ЯВЛЯЕТСЯ ПРЕИМУЩЕСТВОМ:



- 3 Владение английским языком на уровне Intermediate (6.5 IELTS/ 90 TOEFL ibt) и выше



▶ Участие в конкурсе



НЕОБХОДИМЫЙ ПАКЕТ ДОКУМЕНТОВ:

- 1. Заполненная заявка
- 2. Копия удостоверения личности
- 3. Копии дипломов, сертификатов о прохождении курсов (по программированию, статистике, пр.), участия в олимпиадах по математике, информатике или другим техническим предметам
- 4. Копия транскрипта (все семестры) и диплома об окончании бакалавриата с оценками (для магистрантов и специалистов)
- 5. Ответ на вопрос «Я хочу научиться анализировать данные, чтобы...»
- 6. Резюме с портфолио, демонстрирующим достижения кандидата в области IT (при наличии)
- 7. Сертификат, подтверждающий владение английским языком (при наличии)



YESSENOV
DATA LAB



Программа обучения

Неделя 1. Язык программирования Python

11 июня - 15 июня



Куаныш Абешев
декан Школы
инженерного
менеджмента AlmaU



Тимур Бакибаев
Профессор AlmaU

День 1

09:00 – 10:00	Регистрация участников
10:00 – 11:30	Что такое Data Mining, Big Data, примеры
11:30 – 11:45	Кофе-брейк
12:00 – 13:15	Разбор примера Titanic на Kaggle
13:15 – 14:30	Обед
14:30 – 16:00	Введение в Python, переменные, списки, условия, циклы
16:00 – 16:15	Кофе-брейк
16:15 – 18:00	Лабораторная работа: основы Python

День 2

10:00 – 11:30	Структуры данных: списки, множества, библиотеки (ключ-значение)
11:30 – 11:45	Кофе-брейк
12:00 – 13:15	Знакомство с библиотекой NumPy
13:15 – 14:30	Обед
14:30 – 16:00	Лабораторная работа: структуры данных и NumPy
16:00 – 18:00	Team Building

День 3

10:00 – 11:30	Знакомство с библиотекой Pandas, SciPy. Загрузка данных.
11:30 – 11:45	Кофе-брейк
12:00 – 13:15	Группировка данных. Фильтры, сортировка.
13:15 – 14:30	Обед
14:30 – 16:00	Лабораторная работа: работа с CSV, TXT, Quandl.
16:00 – 16:15	Кофе-брейк
16:15 – 18:00	Лабораторная работа: работа с CSV, TXT, Quandl.

День 4

10:00 – 11:30	Объектно-ориентированное программирование
11:30 – 11:45	Кофе-брейк
12:00 – 13:15	Разбор игры Coders Strike Back на codingame.com
13:15 – 14:30	Обед
14:30 – 16:00	Лабораторная работа на codingame.com: Простые задачи
16:00 – 16:15	Кофе-брейк
16:15 – 18:00	Лабораторная работа на codingame.com: Coders Strike Back

День 5

10:00 – 11:30	Загрузка данных. Подготовка данных.
11:30 – 11:45	Кофе-брейк
12:00 – 13:15	Простейшая визуализация (2D Arrays)
13:15 – 14:30	Обед
14:30 – 16:00	Лабораторная работа: Pandas
16:00 – 16:15	Кофе-брейк
16:15 – 18:00	Лабораторная работа: Matplotlib



YESSENOV
DATA LAB



Программа обучения

Неделя 2. Линейные модели классификации и регрессии

18-22 июня



Дмитрий Русанов
Data Scientist,
EPAM Systems

День 1

10:00 – 11:30	Задача оптимизации, метод градиентного спуска
11:30 – 11:45	Кофе-брейк
12:00 – 13:15	Лабораторная работа
13:15 – 14:30	Обед
14:30 – 16:00	Лабораторная работа
16:00 – 16:15	Кофе-брейк
16:15 – 18:00	Лабораторная работа

День 2

10:00 – 11:30	Линейные модели классификации и регрессии
11:30 – 11:45	Кофе-брейк
12:00 – 13:15	Лабораторная работа
13:15 – 14:30	Обед
14:30 – 16:00	Лабораторная работа
16:00 – 16:15	Кофе-брейк
16:15 – 18:00	Лабораторная работа

День 3

10:00 – 11:30	Переобучение, генерализация
11:30 – 11:45	Кофе-брейк
12:00 – 13:15	Лабораторная работа
13:15 – 14:30	Обед
14:30 – 16:00	Лабораторная работа
16:00 – 18:00	Team Building

День 4

10:00 – 11:30	Кросс-валидация
11:30 – 11:45	Кофе-брейк
12:00 – 13:15	Лабораторная работа
13:15 – 14:30	Обед
14:30 – 16:00	Лабораторная работа
16:00 – 16:15	Кофе-брейк
16:15 – 18:00	Лабораторная работа

День 5

10:00 – 11:30	Метрики качества.
11:30 – 11:45	Кофе-брейк
12:00 – 13:15	Лабораторная работа
13:15 – 14:30	Обед
14:30 – 16:00	Лабораторная работа
16:00 – 16:15	Кофе-брейк
16:15 – 18:00	Лабораторная работа





YESSENOV
DATA LAB



Программа обучения

Неделя 3. Работа с признаками (РСА, классификация)

25-29 июня



Михаил Липкович
Lead big data engineer,
EPAM Systems

День 1

10:00 – 11:30	Классификация, деревья решений и метод ближайших соседей
11:30 – 11:45	Кофе-брейк
12:00 – 13:15	Лабораторная работа
13:15 – 14:30	Обед
14:30 – 16:00	Лабораторная работа
16:00 – 16:15	Кофе-брейк
16:15 – 18:00	Лабораторная работа

День 2

10:00 – 11:30	Композиции: бэггинг, бустинг, случайный лес
11:30 – 11:45	Кофе-брейк
12:00 – 13:15	Лабораторная работа
13:15 – 14:30	Обед
14:30 – 16:00	Лабораторная работа
16:00 – 16:15	Кофе-брейк
16:15 – 18:00	Лабораторная работа

День 3

10:00 – 11:30	Обучение без учителя: РСА, кластеризация
11:30 – 11:45	Кофе-брейк
12:00 – 13:15	Лабораторная работа
13:15 – 14:30	Обед
14:30 – 16:00	Лабораторная работа
16:00 – 18:00	Team Building

День 4

10:00 – 11:30	Построение и отбор признаков
11:30 – 11:45	Кофе-брейк
12:00 – 13:15	Лабораторная работа
13:15 – 14:30	Обед
14:30 – 16:00	Лабораторная работа
16:00 – 16:15	Кофе-брейк
16:15 – 18:00	Лабораторная работа

День 5

10:00 – 11:30	Метод опорных векторов (SVM)
11:30 – 11:45	Кофе-брейк
12:00 – 13:15	Лабораторная работа
13:15 – 14:30	Обед
14:30 – 16:00	Лабораторная работа
16:00 – 16:15	Кофе-брейк
16:15 – 18:00	Лабораторная работа



Программа обучения

Неделя 4. Нейронные сети

2-6 июля



Марина Горлова
Analyst,
Yandex Money

День 1

10:00 – 11:30	Введение в нейронные сети. Перцептрон
11:30 – 11:45	Кофе-брейк
12:00 – 13:15	Алгоритм обратного распространения ошибки (Backpropagation)
13:15 – 14:30	Обед
14:30 – 16:00	Лабораторная работа: реализация нейронной сети
16:00 – 16:15	Кофе-брейк
16:15 – 18:00	Лабораторная работа: реализация нейронной сети

День 2

10:00 – 11:30	Введение в библиотеку Keras
11:30 – 11:45	Кофе-брейк
12:00 – 13:15	Введение в библиотеку Keras. Продолжение
13:15 – 14:30	Обед
14:30 – 16:00	Лабораторная работа
16:00 – 16:15	Кофе-брейк
16:15 – 18:00	Лабораторная работа

День 3

10:00 – 11:30	Сверточные нейронные сети
11:30 – 11:45	Кофе-брейк
12:00 – 13:15	Лабораторная работа с анализом изображений
13:15 – 14:30	Обед
14:30 – 16:00	Лабораторная работа с анализом изображений
16:00 – 18:00	Team Building

День 4

10:00 – 11:30	Рекуррентные нейронные сети
11:30 – 11:45	Кофе-брейк
12:00 – 13:15	Лабораторная работа с анализом текстов
13:15 – 14:30	Обед
14:30 – 16:00	Лабораторная работа с анализом текстов
16:00 – 16:15	Кофе-брейк
16:15 – 18:00	Лабораторная работа с анализом текстов

День 5

10:00 – 11:30	Проблема переобучения нейронных сетей, аугментация данных
11:30 – 11:45	Кофе-брейк
12:00 – 13:15	Лабораторная работа
13:15 – 14:30	Обед
14:30 – 16:00	Лабораторная работа
16:00 – 16:15	Кофе-брейк
16:15 – 18:00	Лабораторная работа



YESSENOV
DATA LAB

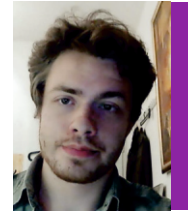


Программа обучения

Неделя 5.

Глубокое обучение в компьютерном зрении
и обучении с подкреплением.
Решение Kaggle задач

9-13 июля



Дмитрий Котовенко
Computer Vision
Research Assistant,
AGT International

День 1

10:00 – 11:30 Классификация MNIST, Fashion MNIST, LFW датасетов
11:30 – 11:45 Кофе-брейк
12:00 – 13:15 Лабораторная работа: Работа над примером
13:15 – 14:30 Обед
14:30 – 16:00 Лабораторная работа: Работа над примером
16:00 – 16:15 Кофе-брейк
16:15 – 18:00 Лабораторная работа: Работа над примером

День 2

10:00 – 11:30 VGG, ResNet и Inception архитектуры. Что видят нейронные сети
11:30 – 11:45 Кофе-брейк
12:00 – 13:15 Лабораторная работа: Работа над примером
13:15 – 14:30 Обед
14:30 – 16:00 Лабораторная работа: Работа над примером
16:00 – 16:15 Кофе-брейк
16:15 – 18:00 Лабораторная работа: Работа над примером

День 3

10:00 – 11:30 От классификации к сегментации. Разбор Kaggle соревнования
11:30 – 11:45 Кофе-брейк
12:00 – 13:15 Лабораторная работа: Работа над примером
13:15 – 14:30 Обед
14:30 – 16:00 Лабораторная работа: Работа над примером
16:00 – 18:00 Team Building

День 4

10:00 – 11:30 Автоэнкодеры и вариационные автоэнкодеры. Pose estimation
11:30 – 11:45 Кофе-брейк
12:00 – 13:15 Лабораторная работа: Работа над примером
13:15 – 14:30 Обед
14:30 – 16:00 Лабораторная работа: Работа над примером
16:00 – 16:15 Кофе-брейк
16:15 – 18:00 Лабораторная работа: Работа над примером

День 5

10:00 – 11:30 Обучение с подкреплением. Пределы supervised learning
11:30 – 11:45 Кофе-брейк
12:00 – 13:15 Лабораторная работа: Работа над примером
13:15 – 14:30 Обед
14:30 – 16:00 Лабораторная работа: Работа над примером
16:00 – 16:15 Кофе-брейк
16:15 – 18:00 Лабораторная работа: Работа над примером



YESSENOV
DATA LAB



Kaspi.kz



Программа обучения

Неделя 6. Лаборатория Kaspi

16-20 июля

День 1 Кто такие аналитики и с чем они работают?

10:00 – 11:30	Кто такой аналитик и зачем он нужен? (Часть 1)
11:30 – 11:45	Кофе - брейк
12:00 – 13:15	Кто такой аналитик и зачем он нужен? (Часть 2)
13:15 – 14:30	Обед
14:30 – 16:00	Практический кейс «Посвящение в аналитики?». Часть 1
16:00 – 16:15	Кофе - брейк
16:15 – 18:00	Практический кейс «Посвящение в аналитики?». Часть 2

День 2 С чего начать?

10:00 – 11:30	Клиентская аналитика – что за «фрукт»?
11:30 – 11:45	Кофе - брейк
12:00 – 13:15	CRM + Аналитика
13:15 – 14:30	Обед
14:30 – 16:00	Развитие ключевых навыков аналитика. Часть 1
16:00 – 16:15	Кофе - брейк
16:15 – 18:00	Развитие ключевых навыков аналитика. Часть 2

День 3 Интеллектуальный Риск

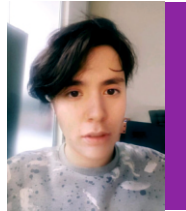
10:00 – 11:30	Быть или не быть кредиту, вот в чем вопрос?
11:30 – 11:45	Кофе - брейк
12:00 – 13:15	«Семь раз отмерь, один раз отрежь».
13:15 – 14:30	Обед
14:30 – 16:00	Поведенческая аналитика, как одна из главных линий защит в антифрод-процессе. Часть 1
16:00 – 16:15	Кофе - брейк
16:15 – 18:00	Поведенческая аналитика, как одна из главных линий защит в антифрод-процессе. Часть 2

День 4 Искусственный интеллект в Kaspi

10:00 – 11:30	Умеешь ли ты читать между строк? Часть 1
11:30 – 11:45	Кофе - брейк
12:00 – 13:15	Умеешь ли ты читать между строк? Часть 2
13:15 – 14:30	Обед
14:30 – 16:00	Когда система знает лучше клиента. Часть 1
16:00 – 16:15	Кофе - брейк
16:15 – 18:00	Когда система знает лучше клиента. Часть 2

День 5 Кейсы из маркетинга

10:00 – 11:30	Что делать, что делать? Однозначно покупать! Часть 1
11:30 – 11:45	Кофе - брейк
12:00 – 13:15	Практический кейс: «Каждому клиенту свой продукт». Часть 1
13:15 – 14:30	Обед
14:30 – 16:00	Практический кейс: «Каждому клиенту свой продукт». Часть 2
16:00 – 16:15	Кофе - брейк
16:15 – 18:00	Практический кейс: «Каждому клиенту свой продукт». Часть 3



Думан Уватаев
Chief Data Officer



Айгерим Сагандыкова
Главный аналитик группы
экспериментальных
проектов



Ильяс Жубанов
Начальник Управления
анализа данных



YESSENOV
DATA LAB



Kaspi.kz



Программа обучения

Неделя 7. Лаборатория Kaspi

23-27 июля

■ Лаборатория Kaspi в цифрах

8000+

студентов прослушали презентацию

7

крупнейших профильных университетов - партнеры

1 500+

студентов пришли на тестирование

100+

студентов успешно сдали тест и прошли обучение

420+

академических часов прослушано

40

прикладных задач решено

16

полноценных аналитических сервисов разработано

90%

студентов нашли отличную работу

■ Студенты Лаборатории Kaspi на основе методов машинного обучения научились:

- ▶ Оценивать риск-профиль клиентов
путем разработки архитектуры автоматической системы принятия решений по принципу «кредитного конвейера»;

- ▶ Оптимизировать операционные процессы
путем централизации контура принятия решений и снижения ресурсоемкости процессов;

- ▶ Разрабатывать справедливые условия оценки любых маркетинговых активностей вне зависимости от каналов коммуникации (массовые, персонализированные);

- ▶ Разрабатывать, внедрять и оценивать различные рекомендательные системы на сайте на основании данных по поведению на сайте;

- ▶ Обособлять первичное от вторичного
при разработке дизайна отчета или контента презентаций / аналитических сводок;

- ▶ Разрабатывать решения компьютерного зрения
обнаружение, сопоставление, отслеживание и классификация объектов;

- ▶ Понимать бизнес
и имплементировать data driven процессы в компании.



YESSENOV
DATA LAB



Программа обучения

Неделя 8. Проектные задачи

30 июля - 3 августа

Казахстанские компании, использующие анализ данных в своей работе, представят перед обучающимися реальные задачи, которые актуальны на сегодняшний день. Успешные выпускники школы получают предложения о работе (job offer).



ChocoFamily



Chocolife[®] | Chocotravel | Chocofood | Lensmark[®] | Chocomart[®]





YESSENOV
DATA LAB

ПРОГРАММЫ ФОНДА

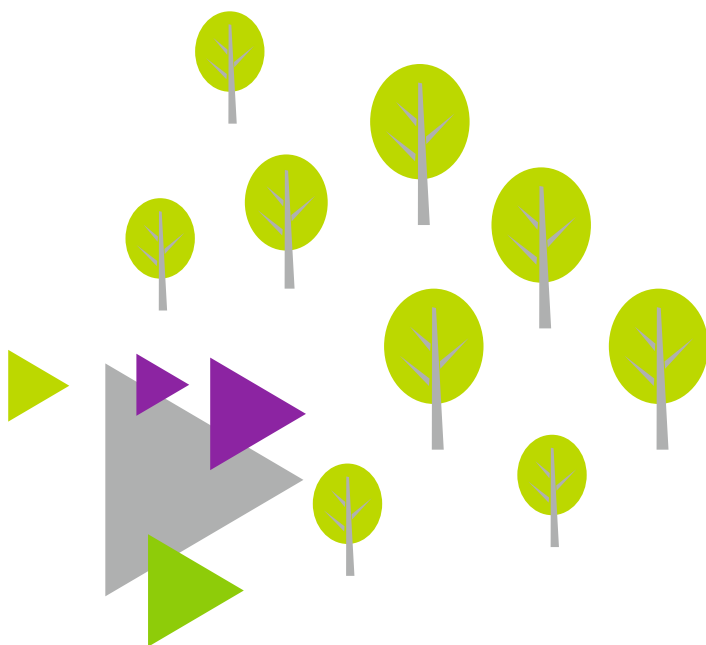
Знания	Наука	Ресурсы
 Есеновские чтения	 Научные стажировки	 Развитие IT компетенций
 Английский язык	 Стипендия Есенова	 Команда SOS
 Популяризация науки	 Обучение в магистратуре	 Повышение личной эффективности
 Казахстанская федерация шахмат	 Алматы марафон	 Алматинская федерация триатлона

Мы: yessenovfoundation.org

ВСЕГДА: info@yessenovfoundation.org

НА СВЯЗИ: +7 727 346 92 88





Алматы, 2018

В партнерстве с

