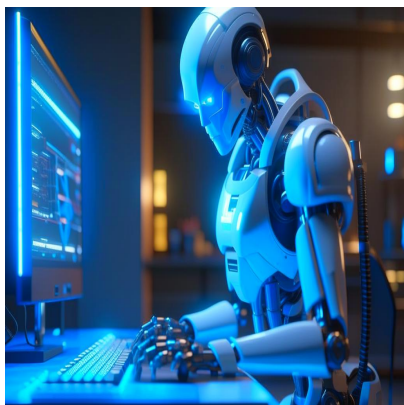




11 лучших ресурсов по науке о данных и машинному обучению

Описание

Данные – это новая нефть. А машинное обучение – это огонь. Тот, кто контролирует эти два направления, будет контролировать весь мир. Нет, все вышесказанное – не напыщенная фраза, взятая из антиутопического романа. **Это реальность.** Новый мировой порядок – это сбор огромного количества данных и их обработка в практические выводы, чего человечество еще не умело делать за всю свою историю. Именно такие технологии позволяют стране опередить других и, в конечном счете, править миром. В результате прогрессивные страны мира относятся к нему очень и очень серьезно.

Прибыльный выбор профессии

Если отбросить международные интриги, наука о данных и машинное обучение – это новая горячая область с невероятными возможностями. Спрос на них зашкаливает (мягко говоря), а специалистов по изучению данных не хватает. Даже посредственных. Как будто мы внезапно открыли множество новых пригодных для жизни планет, а людей для их переселения не хватает. Я мог бы продолжать и продолжать и звучать как заезженная пластинка, но, по-моему, эта инфографика делает свою работу гораздо лучше: Итак, мы видим, что зарплаты начинаются от \$50 000+, а для менеджеров могут перевалить за \$250 000.

И не только это: средний человек на нашей планете будет генерировать 1,7 МБ данных в секунду. Это 3 500+ ТБ данных за всю жизнь – больше, чем мы можем обработать на данный момент, не говоря уже о том, чтобы использовать для

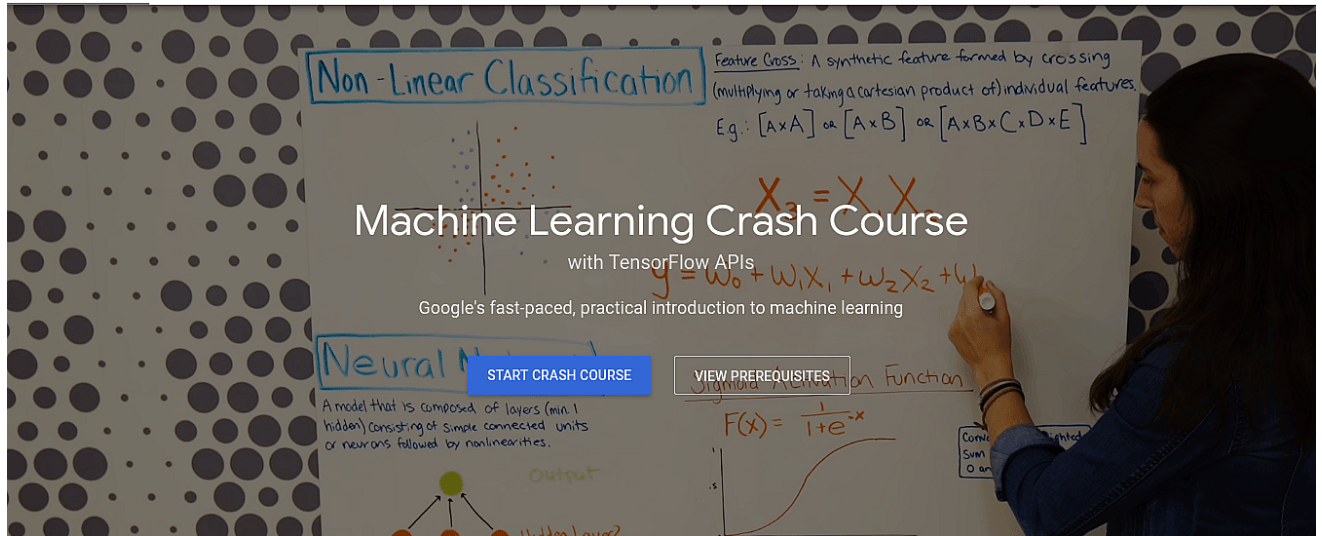
анализа.

Трудно ли заниматься наукой о данных и машинным обучением?

Хороший вопрос! По моему опыту, ответ может быть как “да”, так и “нет”. Искусственный интеллект (и, как следствие, машинное обучение) – сложнейшее занятие, если вы хотите заниматься исследованиями и расширять границы возможностей. Для такой работы недостаточно даже докторской степени по информатике и математике. Но у обычного человека нет ни амбиций, ни времени на такую работу. На другом конце находится то, что я бы назвал прикладной наукой о данных и машинным обучением. То есть вы берете существующие инструменты, методы и алгоритмы и *применяете* их для решения реальных задач. Эта часть требует самоотдачи, восприятия и творческого мышления (а также знания некоторых простых математических понятий, которые быстро усваиваются), но в отношении настоящих “технических” знаний она гораздо более мягкая, чем та, к которой призывает работа инженера-программиста. Другими словами, это не простая прогулка, но по **соотношению вознаграждения и усилий** это одна из лучших инвестиций. Теперь, когда вы твердо решили стать специалистом по изучению данных и инженером машинного обучения, давайте начнем изучать лучшие варианты.

Машинное обучение (Google)

Не многие знают, что в Google есть обширный, очень практичный и бесплатный курс по машинному обучению. По словам представителей компании, это часть их обязательств по продвижению технологий AI/ML и сохранению знаний в открытом доступе.



Самое приятное в этом курсе то, что для него нет предварительных требований, но все же подготовьтесь к тому, что вам придется потратить дополнительное время на самостоятельное изучение концепций статистики. Я имею в виду, что это не обязательно, но если у вас нет опыта в продвинутой статистике, объяснений в этом курсе может быть недостаточно. Еще одна загвоздка заключается в том, что этот курс знакомит с машинным обучением через TensorFlow, который является реализацией ML, разработанной Google. Таким образом, Google в некотором смысле стремится продвинуть свои API для машинного обучения, но, учитывая ценность данного курса, я не вижу, как это может стать камнем преткновения. Если что, TensorFlow – один из самых простых способов освоить ML и пользуется бешеной популярностью.

Наука о данных

Название **Гарвард** внушает благоговение, так же как и этот курс. Прежде всего, это не курс “Давайте быстро разберемся с грязью”, в котором вы будете ходить на цыпочках вокруг Machine Learning, написав сниппет здесь или скрипт здесь. Этот курс – суровое боевое крещение, требующее упорной работы и значительных затрат времени. К курсу прилагаются бесплатные видео, код (размещенный на GitHub) и решения лабораторных заданий, так что практически вы ничем не ограничены, если хотите пройти его. Идеальная аудитория? Ты... Я не шучу. Я бы сказал, работающим профессионалам с приличным математическим образованием, даже если они, возможно, уже не увлекаются математикой (привычка к умозаключениям и доказательствам – самое необходимое). Но еще раз

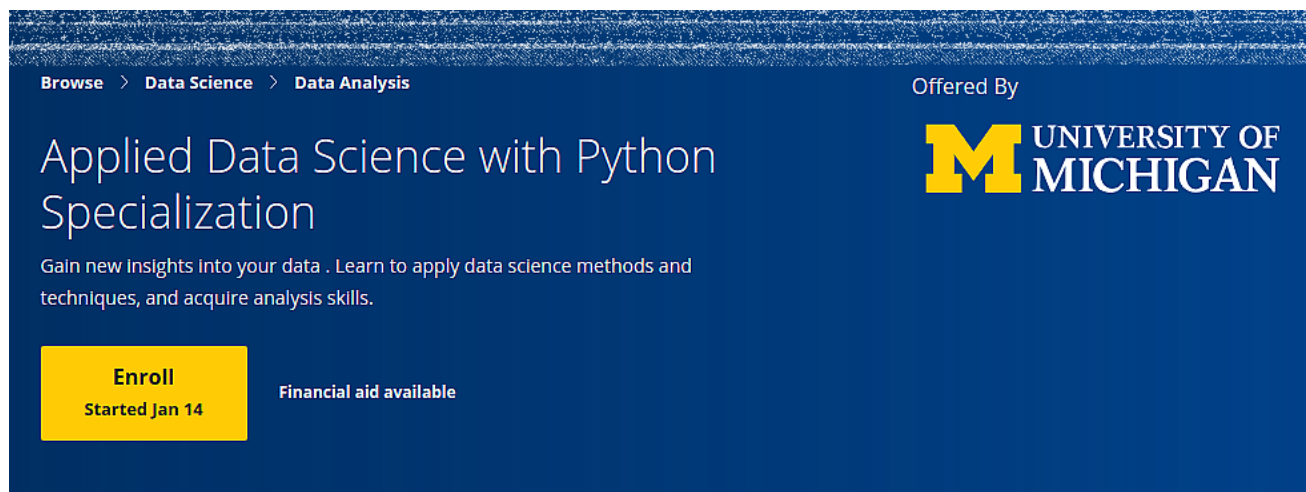
предупреждаю: вы можете думать, что у вас все хорошо, но этот курс покажется вам чем-то вроде завтрака из закаленных гвоздей – практические задачи достаточно сложны, чтобы заставить вас плакать, но, возможно, это как раз то, что вы ищете!

Машинное обучение

Зайдите в бар, где полно специалистов по обработке данных, и спросите, кто такой **Эндрю Нг**, и вы получите отпор. В кругах специалистов по науке о данных и машинному обучению Эндрю Нг приобрел статус бога благодаря своему исключительному курсу на Coursera – Machine Learning. А если вы сомневаетесь в авторитете Эндрю Нг, позвольте мне сказать об этом самому: Это платный курс, поскольку он входит в тарифный план Coursera, но финансовые обязательства и решимость – не единственные необходимые условия. Это длинный курс, поскольку Эндрю глубоко погружается в математику, лежащую в основе всех вещей ML, и разбирает популярные алгоритмы. Но, к счастью, это полный курс, и вас шаг за шагом проведут в самые темные глубины и выведут обратно. Очень рекомендую, главным образом потому, что сегодня стало принято щеголять сертификатом об окончании этого курса!

Прикладная наука о данных

Специализации на Coursera состоят из серии курсов, которые призваны помочь вам пройти путь от нуля до профессионала в определенной концепции. Если вы ищете полный, серьезный, но дружелюбный курс по Data Science и Machine Learning с Python, я не могу не порекомендовать вам эту специализацию.



Browse > Data Science > Data Analysis

Offered By

M UNIVERSITY OF MICHIGAN

Applied Data Science with Python Specialization

Gain new insights into your data. Learn to apply data science methods and techniques, and acquire analysis skills.

Enroll
Started Jan 14

Financial aid available

По окончании курса вы получаете сертификат.

DataCamp

DataCamp предлагает множество курсов по науке о данных, которые также включают в себя несколько навыков и карьерных направлений. От манипулирования данными до машинного обучения – вы получите навыки карьерного роста в Python и R, которые помогут вам добиться успеха в области науки о данных.

The image is a screenshot of the DataCamp website's main landing page. It features a dark blue background with white and yellow text. At the top left is the DataCamp logo, which consists of a blue hexagon with a white head icon inside, followed by the word "DataCamp" in white. To the right of the logo is a search bar with a magnifying glass icon and the word "Search". Below the logo and search bar, the text "THE SMARTEST WAY TO" is written in yellow, followed by "Learn Data Science Online" in large white letters. Underneath this, a paragraph of white text explains the value of data science skills. Two buttons are present: a yellow "Start Learning For Free" button and a white "DataCamp For Enterprise" button. At the bottom, there is a row of logos for various technologies: Python, R, SQL, Spreadsheets, Git, Shell, Scala, Apache Spark, Tableau, Oracle, and Power BI.

 **DataCamp**

THE SMARTEST WAY TO

Learn Data Science Online

The skills people and businesses need to succeed are changing. No matter where you are in your career or what field you work in, you will need to understand the language of data. With DataCamp, you learn data science today and apply it tomorrow.

[Start Learning For Free](#) [DataCamp For Enterprise](#)

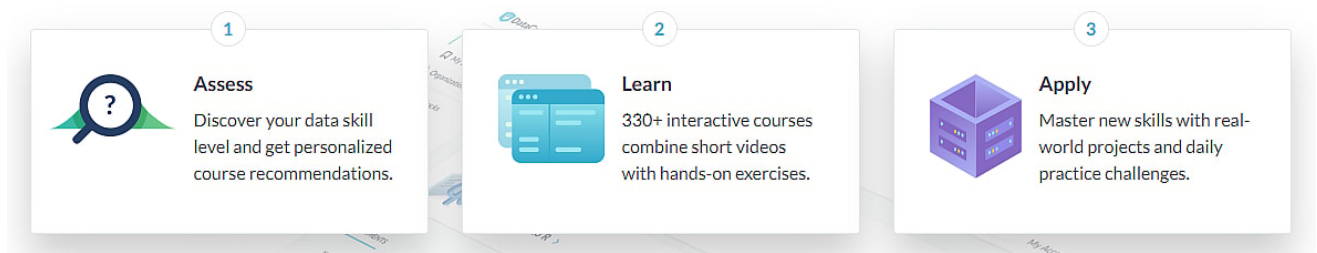
 python™  R  SQL  SPREADSHEETS

 git  Shell  Scala   X

 + a b | e a u  ORACLE™  Power BI

Благодаря байтовому контенту DataCamp вы можете учиться в своем собственном темпе. На этих курсах вы получите практический опыт, который поможет вам усовершенствовать свои навыки работы с данными.

Learn The Way You Learn Best



Вы можете начать с бесплатной версии и оценить курс, ознакомившись с первой главой.

edX

Учитесь у MITx, Harvardx, IBM, RICEx, UCSanDiegox и GTx на платформе edX. Все они имеют обширную учебную программу, которая поможет вам приобрести навыки специалиста по работе с данными. Эти программы лучше всего подходят для тех, кто занимается статистикой или компьютерными науками.

Data science courses on edX

Data science is a multifaceted field used to gain insights from complex data.

Data scientists use knowledge of

- [Probability and Statistics](#) ⬇
- [Excel and Business Analytics](#) ⬇
- [Python](#) ⬇
- [R](#) ⬇

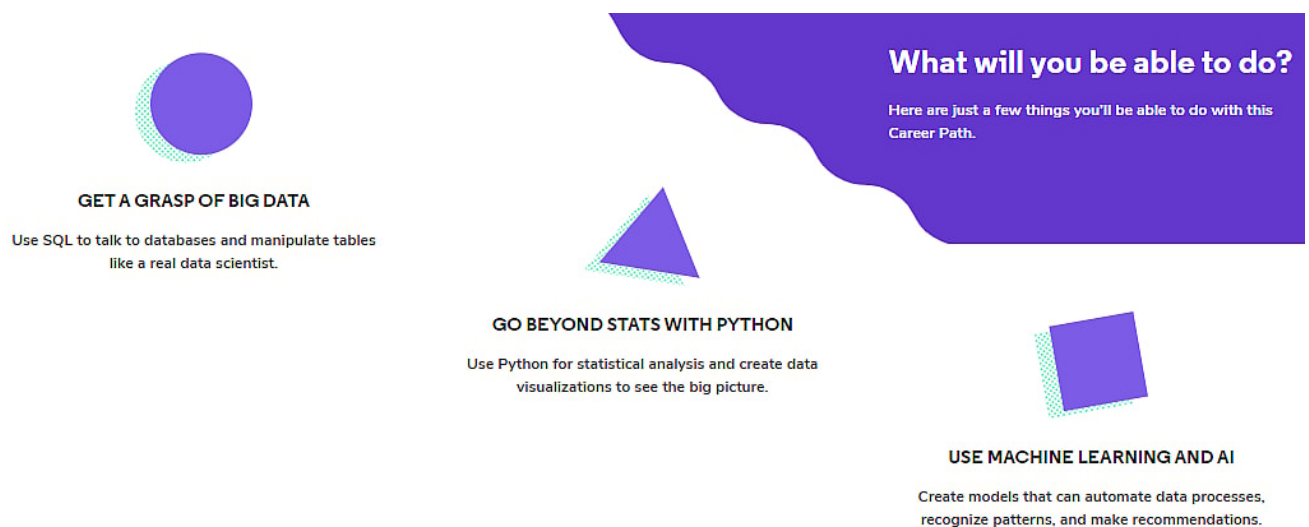
What is edX?

Real college courses from Harvard, MIT, and more of the world's leading universities.

Если вам не нужна программа, вы можете выбрать аля-карт. На сайте edX вы найдете более 200 курсов, связанных с наукой о данных, которые охватывают Python, R, Excel, вероятность, статистику, машинное обучение, визуализацию данных и многое другое.

Codecademy

Codecademy – еще одна платформа, которая является одной из лучших систем, помогающих научиться кодить. Они верят в принцип “Учись на практике” и предлагают множество практических проектов и тестов на своей платформе. Курс по науке о данных, предлагаемый Codecademy, включает в себя SQL, Python 3, NumPy, Pandas, Matplotlib, Scikit-learn и многие другие библиотеки.



Вся карьерная программа включает в себя 26 курсов, которых более чем достаточно, чтобы стать успешным специалистом по изучению данных. Этот курс данных:

- Дает глубокие знания в области науки о данных
- Обеспечивает легкое следование дорожной карте
- Поможет вам подготовиться к работе и приобрести достаточный практический опыт

Udemy

Udemy не нуждается в представлении. Python for Data Science and Machine Learning Bootcamp на Udemy – один из самых популярных курсов с рейтингом более 85 тысяч человек (4,6), который изучают 370 тысяч студентов по всему миру.

Development > Data Science > Machine Learning

Python for Data Science and Machine Learning Bootcamp

Learn how to use NumPy, Pandas, Seaborn , Matplotlib , Plotly , Scikit-Learn , Machine Learning, Tensorflow , and more!

4.6 ★★★★★ (84,872 ratings) 377,520 students

Ниже перечислены темы, рассматриваемые в этом курсе:

What you'll learn

- | | |
|--|---|
| ✓ Use Python for Data Science and Machine Learning | ✓ Use Spark for Big Data Analysis |
| ✓ Implement Machine Learning Algorithms | ✓ Learn to use NumPy for Numerical Data |
| ✓ Learn to use Pandas for Data Analysis | ✓ Learn to use Matplotlib for Python Plotting |
| ✓ Learn to use Seaborn for statistical plots | ✓ Use Plotly for interactive dynamic visualizations |
| ✓ Use SciKit-Learn for Machine Learning Tasks | ✓ K-Means Clustering |
| ✓ Logistic Regression | ✓ Linear Regression |
| ✓ Random Forest and Decision Trees | ✓ Natural Language Processing and Spam Filters |
| ✓ Neural Networks | ✓ Support Vector Machines |

Ниже перечислены особенности/возможности этого курса:

- 25 часов видео по запросу
- Полный пожизненный доступ
- 13 статей и пять ресурсов для скачивания
- Доступ с мобильных и телевизионных устройств
- Сертификат о прохождении
- 30-дневная гарантия возврата денег

Поэтому, если вы предпочитаете бюджетные курсы, этот подойдет вам для начала.

Google AI

Интересно ли вам будет изучать машинное обучение у экспертов по ML в Google? Тогда вам необходимо ознакомиться с курсами по Google AI. На этой платформе есть курсы и материалы по машинному обучению и науке о данных для студентов, инженеров-программистов, ученых, изучающих данные, и даже исследователей. Эти курсы бесплатны. Начнем с того, что вам следует пройти курс Machine Learning Crash Course в Google AI. Это быстро развивающийся курс с практическим введением с использованием API TensorFlow. Ниже приведены подробности этого курса:



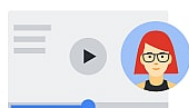
30+ exercises



25 lessons



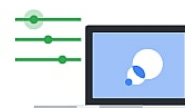
15 hours



Lectures from Google researchers



Real-world case studies



Interactive visualizations of algorithms in action

На этой платформе также есть специальные курсы по таким важным темам машинного обучения, как кластеризация, рекомендательные системы, тестирование и отладка в машинном обучении, разделение данных и разработка функций в машинном обучении. Если вы уже знаете основы машинного обучения, эти курсы будут вам полезны.

Udacity

Udacity – это также очень популярная платформа электронного обучения, которая предлагает множество курсов по современным технологиям. Она имеет несколько ведущих отраслевых программ, разработанных и признанных ведущими компаниями по всему миру, такими как AT&T, AWS, Google, IBM. Одна из программ в Udacity – Data Science – School of Data Science. Эта программа поможет вам получить работу аналитика данных, ученого по данным, инженера по данным и бизнес-аналитика. Курс Data Scientist в этой программе является ключевым и охватывает

концепции машинного обучения, глубокого обучения и программной инженерии. Чтобы выбрать этот курс, необходимо обладать базовыми знаниями в области машинного обучения.

SYLLABUS

Data Scientist

You'll master the skills necessary to become a successful Data Scientist. You'll work on projects designed by industry experts, and learn to run data pipelines, design experiments, build recommendation systems, and deploy solutions to the cloud.

If you already have experience with machine learning, take this program.

[+ SEE DETAILS](#)

Estimated 4 months to complete



Если вы знаете программирование на python, но не знакомы с машинным обучением, на Udacity есть еще одна программа – School of AI. В этой программе есть курсы, начинающиеся с основ машинного обучения.

SYLLABUS

Intro to Machine Learning with TensorFlow

Learn foundational machine learning algorithms, starting with data cleaning and supervised models. Then, move on to exploring deep and unsupervised learning. At each step, get practical experience by applying your skills to code exercises and projects.

This program is intended for students with experience in Python, who have not yet studied Machine Learning topics.

[+ SEE DETAILS](#)

3 months to complete



Глубокое обучение

Этот курс – просто благословение и моя самая любимая рекомендация в этом списке, если вы кодер. Повторю еще раз: если вы кодер. Это потому, что данный курс не уделяет времени обучению основам программирования. В описании курса об этом сказано очень ясно (выделение оригинальное):

Мы предполагаем, что каждый слушатель этого курса имеет **опыт кодирования не менее одного года**. В качестве языка обучения в курсе используется python, поэтому если вы еще не знаете python, мы

предполагаем, что вы потратите время на его изучение – для опытных кодеров python будет довольно простым языком для изучения.



Так что если вы уже знаете Python или можете быстро освоить, то это идеальный курс для прагматиков, которые хотят создавать реальные, удобные системы, не слишком беспокоясь о теоретических основах алгоритмов. Я бы даже сказал, что это для нетерпеливых любителей пошалить (вроде меня!), которые ненавидят церемонии и однообразие. А я уже упоминал, что он на 100% бесплатный и имеет отличное сообщество?!

Заключение

Фух! Составить этот список было очень сложно. Не потому, что не было достаточно хороших источников, а потому, что их было слишком много! Машинное обучение – это область, которая буквально взорвалась и решает сложные проблемы очень элегантно, поэтому в сети существуют сотни курсов, бесплатных и платных, и большинство из них очень, очень хороши. Но это также может стать источником путаницы, поэтому я постарался свести их к одиннадцати для разных типов обучающихся в зависимости от их уровня опыта.

Дата Создания

22.02.2024