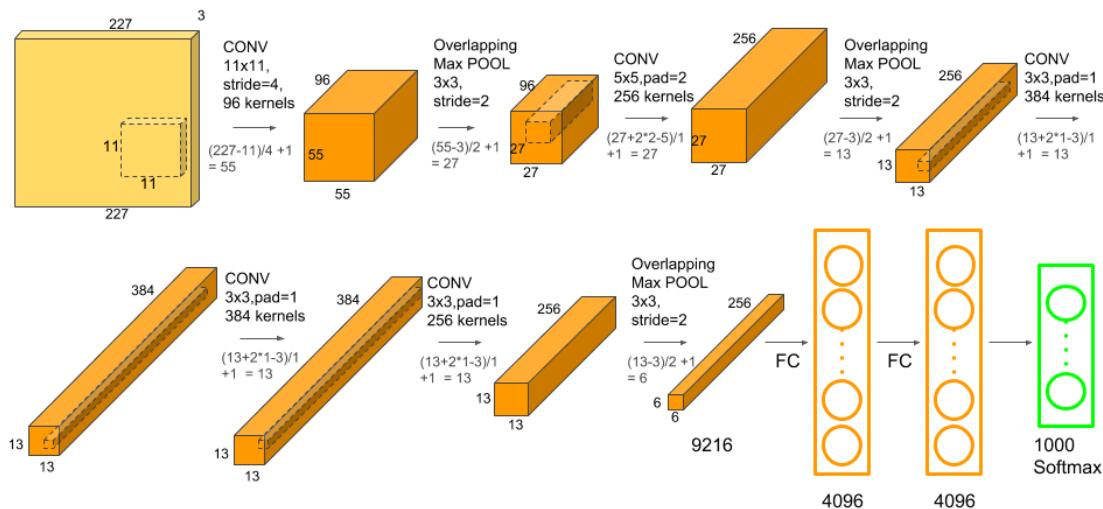


AlexNet – свёрточная нейронная сеть для классификации изображений

1 ноября 2018 [Виды нейросетей](#)



AlexNet – свёрточная нейронная сеть, которая оказалась большое влияние на развитие машинного обучения, в особенности – на алгоритмы компьютерного зрения. Сеть с большим отрывом выиграла конкурс по распознаванию изображений ImageNet LSVRC-2012 в 2012 году (с количеством ошибок 15,3% против 26,2% у второго места).

Архитектура AlexNet схожа с созданной Yann LeCun сетью [LeNet](#). Однако у AlexNet больше фильтров на слое и вложенных свёрточных слоев. Сеть включает в себя свертки, максимальное объединение, дропаут,

аугментацию данных, функции активаций ReLU и стохастический градиентный спуск.

Особенности AlexNet

1. Как функция активации используется Relu вместо арктангенса для добавления в модель нелинейности. За счет этого при одинаковой точности метода скорость становится в 6 раз быстрее.
2. Использование дропаута вместо регуляризации решает проблему переобучения. Однако время обучения удваивается с показателем дропаута 0,5.
3. Производится перекрытие объединений для уменьшения размера сети. За счет этого уровень ошибок первого и пятого уровней снижаются до 0,4% и 0,3%, соответственно.

Датасет ImageNet

ImageNet – набор из 15 миллионов помеченных изображений с высоким разрешением, разделенных на 22 000 категорий. Изображения собраны в интернете и помечены вручную с помощью краудсорсинга Amazon's Mechanical Turk. Начиная с 2010 года проводится ежегодный конкурс ImageNet Large-Scale Visual Recognition Challenge (ILSVRC), являющийся частью Pascal Visual Object Challenge. В челлендже используется часть датасета ImageNet с 1000 изображений в каждой из 1000 категорий. Всего получается 1,2 миллиона изображений для обучения, 50 000 изображений для проверки и 150 000 – для тестирования. ImageNet состоит из изображений с разным разрешением. Поэтому для

конкурса их масштабируют до фиксированного разрешения 256×256 . Если изначально изображение было прямоугольным, то его обрезают до квадрата в центре изображения.

Архитектура

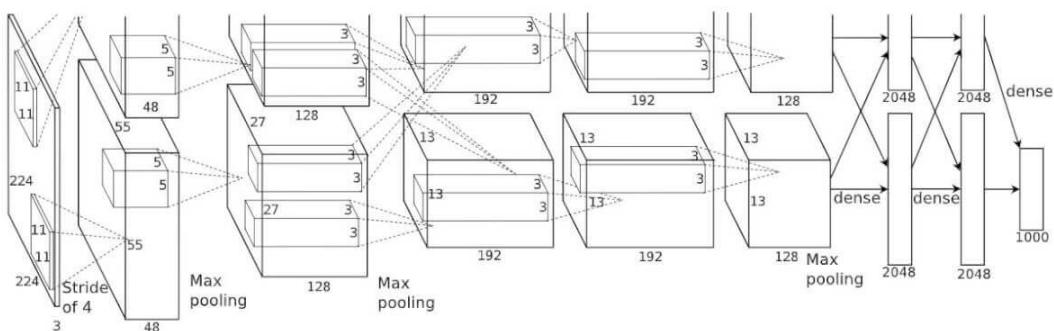


Рисунок 1

Архитектура сети приведена на рисунке 1. AlexNet содержит восемь слоев с весовыми коэффициентами. Первые пять из них сверточные, а остальные три – полносвязные. Выходные данные пропускаются через функцию потерь softmax, которая формирует распределение 1000 меток классов. Сеть максимизирует многолинейную логистическую регрессию, что эквивалентно максимизации среднего по всем обучающим случаям логарифма вероятности правильной маркировки по распределению ожидания. Ядра второго, четвертого и пятого сверточных слоев связаны только с теми картами ядра в предыдущем слое, которые находятся на одном и том же графическом процессоре. Ядра третьего сверточного слоя связаны со всеми картами ядер второго слоя. Нейроны в полносвязных слоях связаны со всеми нейронами предыдущего слоя.

Таким образом, AlexNet содержит 5 сверточных слоев и 3 полносвязных слоя. Relu применяется после каждого

сверточного и полносвязного слоя. Дропаут применяется перед первым и вторым полносвязными слоями. Сеть содержит 62,3 миллиона параметров и затрачивает 1,1 миллиарда вычислений при прямом проходе.

Сверточные слои, на которые приходится 6% всех параметров, производят 95% вычислений.

Обучение

AlexNet проходит 90 эпох. Обучение занимает 6 дней одновременно на двух графических процессорах Nvidia Geforce GTX 580, что является причиной того, что сеть разделена на две части. Используется стохастический градиентный спуск со скоростью обучения 0,01, импульсом 0,9 и распадом весовых коэффициентов 0,0005. Скорость обучения делится на 10 после насыщения точности и снижается в 3 раза в течение обучения. Схема обновления весовых коэффициентов w имеет вид:

$$\begin{aligned} v_{i+1} &:= 0.9 \cdot v_i - 0.0005 \cdot \epsilon \cdot w_i - \epsilon \cdot \left\langle \frac{\partial L}{\partial w} \Big|_{w_i} \right\rangle_{D_i} \\ w_{i+1} &:= w_i + v_{i+1} \end{aligned}$$

где i — номер итерации, v — переменная импульса, а ϵ — скорость обучения. В ходе всего этапа обучения скорость обучения выбиралась равной для всех слоев и корректировалась вручную. Последующая эвристика заключалась в том, чтобы разделить скорость обучения на 10, когда количество ошибок при проверке переставало уменьшаться.

Примеры использования и реализация

Результаты показывают, что большая, глубокая сверточная нейронная сеть способна достигать рекордных результатов на очень сложных датасетах, используя только обучение с учителем. Через год после публикации AlexNet все участники конкурса ImageNet стали использовать сверточные нейронные сети для решения задачи классификации. AlexNet была первой реализацией сверточных нейронных сетей и открыла новую эру исследований. Сейчас реализовать AlexNet стало проще с помощью библиотек глубокого обучения: [PyTorch](#), [TensorFlow](#), [Keras](#).

Результат

Сеть достигает следующего уровня ошибок первого и пятого уровней: 37,5% и 17,0%, соответственно. Лучшая производительность, достигнутая в ходе конкурса ILSVRC-2010, составляла 47,1% и 28,2% при использовании подхода, в котором усредняются предсказания, полученные шестью моделями с разреженным кодированием, обученных на различных векторах свойств. С тех пор достигнуты результаты: 45,7% и 25,7% при использовании подхода, в котором усредняются предсказания двух классификаторов, обучаемых на векторах Фишера. Результаты ILSVRC-2010 приведены в таблице 1.

Model	Top-1	Top-5
<i>Sparse coding [2]</i>	47.1%	28.2%
<i>SIFT + FVs [24]</i>	45.7%	25.7%
CNN	37.5 %	17.0 %

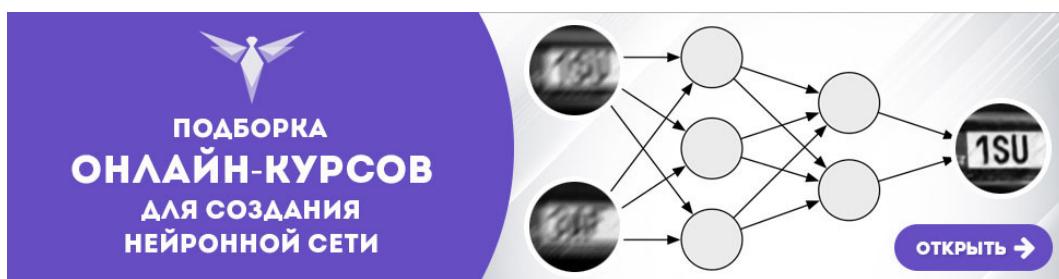


Слева: восемь тестовых изображений ILSVRC-2010 и пять ярлыков, которые наиболее вероятны по мнению модели. Правильная метка записывается под каждым изображением, а вероятность показана красной полосой, если она находится в верхней пятерке. Справа: пять тестовых изображений ILSVRC-2010 в первом столбце. В остальных столбцах показаны шесть обучающих изображений.



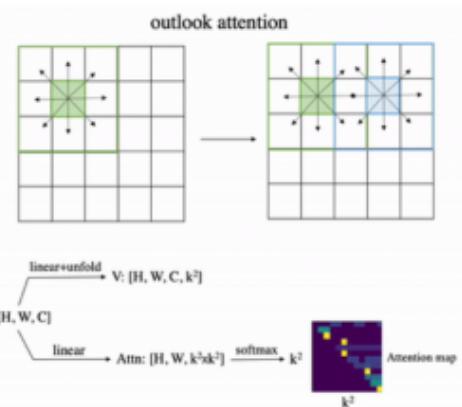
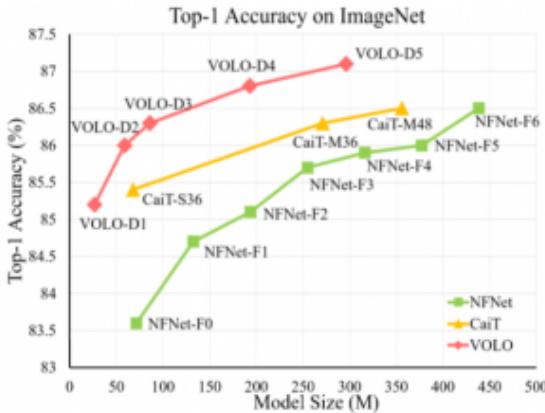
Автор: [Павел Глек](#)

Источник: <https://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks>



Больше интересного

Архитектура Vision Outlooker установила рекорд точности классификации изображений без предобучения



SEER: self-supervised нейросеть с миллиардом параметров от FAIR



ImageNet-A: как снизить точность предсказаний нейросети на 90%

Fox Squirrel



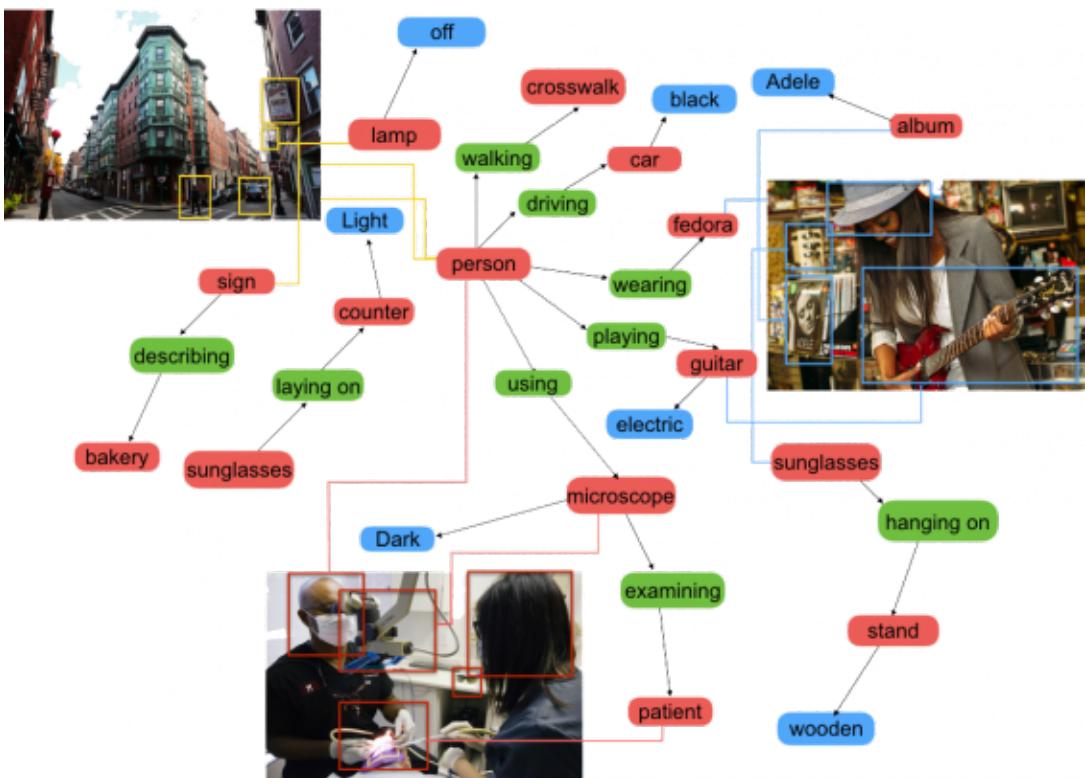
Sea Lion (99%)

Dragonfly

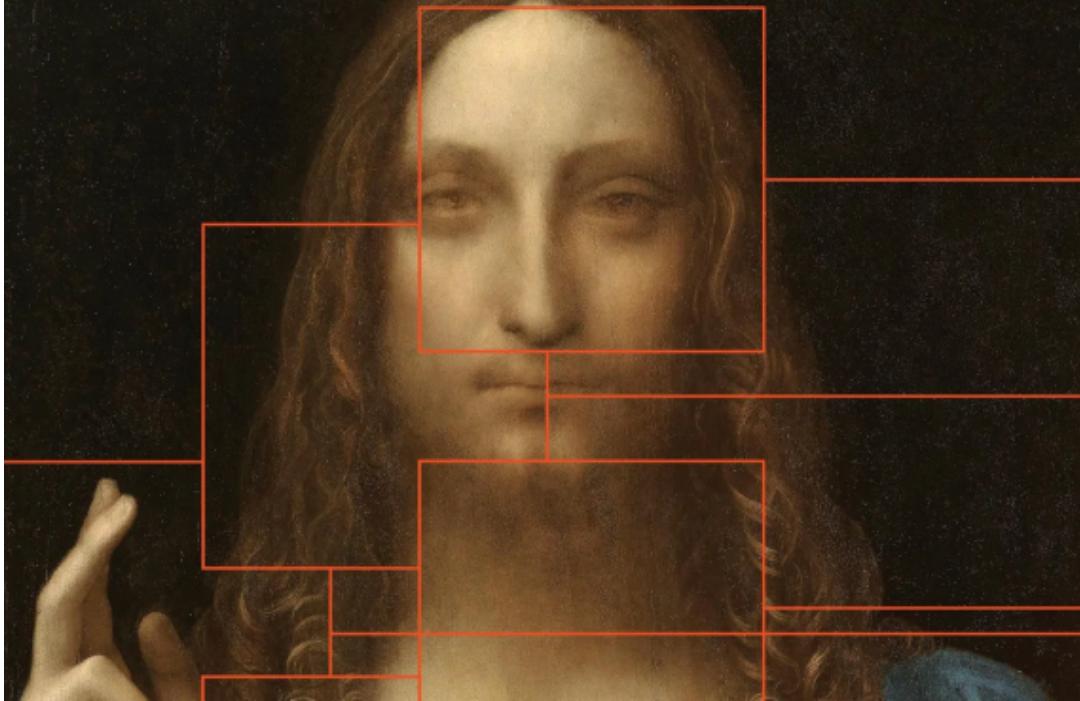
Manhole Cover (99%)



Visual Genome: датасет размеченных изображений



Нейросеть подтвердила подлинность картины да Винчи



Know Your Data: инструмент анализа датасетов Google

Know Your Data coco_captions ⓘ DATASETS STATS RELATIONS ITEM Add Filter

Feature 1	Feature 2	Search rows	Active row	Active column	Actions																																																																																
label	captions_word... cloud_vision label	Row label	-	-																																																																																	
cloud_vision label	coco_captions capti...																																																																																				
<table border="1"> <thead> <tr> <th>LABEL / CAPTIONS_WORDS_GENDERED</th> <th>boy 4,431</th> <th>female 1,029</th> <th>girl 4,057</th> <th>male 1,732</th> <th>man 22,997</th> <th>men 4,741</th> <th>trans 2</th> <th>woman 11,188</th> <th>women 2,583</th> </tr> </thead> <tbody> <tr> <td> Baseball park 1,061</td> <td>↗ 1.39x 170 (122)</td> <td>↘ 4.73x 6 (28.4)</td> <td>↘ 10.18x 11 (112)</td> <td>↘ 1.04x 46 (47.8)</td> <td>↗ 1.22x 777 (635)</td> <td>↗ 1.57x 206 (131)</td> <td>↘ 0 (0.1)</td> <td>↘ 30.87x 10 (309)</td> <td>↘ 17.82x 4 (71.3)</td> </tr> <tr> <td> Toddler 691</td> <td>↗ 5.06x 403 (79.6)</td> <td>↘ 6.16x 3 (18.5)</td> <td>↗ 3.8x 277 (72.9)</td> <td>↘ 4.45x 7 (31.1)</td> <td>↘ 7.13x 58 (413)</td> <td>↘ 85.2x 1 (85.2)</td> <td>↘ 0 (0)</td> <td>↘ 2.45x 82 (201)</td> <td>↘ 11.61x 4 (46.4)</td> </tr> <tr> <td> Batting glove 639</td> <td>↗ 1.64x 121 (73.6)</td> <td>↘ 8.55x 2 (17.1)</td> <td>↘ 13.48x 5 (67.4)</td> <td>↗ 1.01x 29 (28.8)</td> <td>↗ 1.3x 496 (382)</td> <td>↗ 1.12x 88 (78.8)</td> <td>↘ 0 (0)</td> <td>↘ 26.56x 7 (186)</td> <td>↘ 14.31x 3 (42.9)</td> </tr> <tr> <td> Kickflip 635</td> <td>↗ 3.13x 229 (73.2)</td> <td>↘ 16.99x 1 (17)</td> <td>↘ 5.15x 13 (67)</td> <td>↗ 2.31x 66 (28.6)</td> <td>↗ 1.51x 573 (380)</td> <td>↘ 3.01x 26 (78.3)</td> <td>↘ 0 (0)</td> <td>↘ 26.4x 7 (185)</td> <td>↘ 42.66x 1 (42.7)</td> </tr> <tr> <td> Baseball 1,246</td> <td>↗ 1.81x 260 (144)</td> <td>↘ 5.56x 6 (33.3)</td> <td>↘ 5.48x 24 (131)</td> <td>↘ 1.08x 52 (56.1)</td> <td>↗ 1.18x 876 (745)</td> <td>↗ 1.48x 228 (154)</td> <td>↘ 0 (0.1)</td> <td>↘ 19.08x 19 (363)</td> <td>↘ 9.3x 9 (83.7)</td> </tr> <tr> <td> Skateboard deck 660</td> <td>↗ 2.73x 208 (76.1)</td> <td>↘ 8.83x 2 (17.7)</td> <td>↘ 6.33x 11 (69.6)</td> <td>↗ 2.35x 70 (29.7)</td> <td>↗ 1.51x 598 (395)</td> <td>↘ 3.26x 25 (81.4)</td> <td>↘ 0 (0)</td> <td>↘ 17.46x 11 (192)</td> <td>↘ 0 (44.3)</td> </tr> <tr> <td> Baseball protective gear 491</td> <td>↗ 2x 113 (56.6)</td> <td>↘ 4.38x 3 (13.1)</td> <td>↘ 4.32x 12 (51.8)</td> <td>↘ 1.01x 22 (22.1)</td> <td>↗ 1.13x 331 (294)</td> <td>↗ 1.47x 89 (60.5)</td> <td>↘ 0 (0)</td> <td>↘ 20.41x 7 (143)</td> <td>↘ 16.49x 2 (33)</td> </tr> </tbody> </table>						LABEL / CAPTIONS_WORDS_GENDERED	boy 4,431	female 1,029	girl 4,057	male 1,732	man 22,997	men 4,741	trans 2	woman 11,188	women 2,583	Baseball park 1,061	↗ 1.39x 170 (122)	↘ 4.73x 6 (28.4)	↘ 10.18x 11 (112)	↘ 1.04x 46 (47.8)	↗ 1.22x 777 (635)	↗ 1.57x 206 (131)	↘ 0 (0.1)	↘ 30.87x 10 (309)	↘ 17.82x 4 (71.3)	Toddler 691	↗ 5.06x 403 (79.6)	↘ 6.16x 3 (18.5)	↗ 3.8x 277 (72.9)	↘ 4.45x 7 (31.1)	↘ 7.13x 58 (413)	↘ 85.2x 1 (85.2)	↘ 0 (0)	↘ 2.45x 82 (201)	↘ 11.61x 4 (46.4)	Batting glove 639	↗ 1.64x 121 (73.6)	↘ 8.55x 2 (17.1)	↘ 13.48x 5 (67.4)	↗ 1.01x 29 (28.8)	↗ 1.3x 496 (382)	↗ 1.12x 88 (78.8)	↘ 0 (0)	↘ 26.56x 7 (186)	↘ 14.31x 3 (42.9)	Kickflip 635	↗ 3.13x 229 (73.2)	↘ 16.99x 1 (17)	↘ 5.15x 13 (67)	↗ 2.31x 66 (28.6)	↗ 1.51x 573 (380)	↘ 3.01x 26 (78.3)	↘ 0 (0)	↘ 26.4x 7 (185)	↘ 42.66x 1 (42.7)	Baseball 1,246	↗ 1.81x 260 (144)	↘ 5.56x 6 (33.3)	↘ 5.48x 24 (131)	↘ 1.08x 52 (56.1)	↗ 1.18x 876 (745)	↗ 1.48x 228 (154)	↘ 0 (0.1)	↘ 19.08x 19 (363)	↘ 9.3x 9 (83.7)	Skateboard deck 660	↗ 2.73x 208 (76.1)	↘ 8.83x 2 (17.7)	↘ 6.33x 11 (69.6)	↗ 2.35x 70 (29.7)	↗ 1.51x 598 (395)	↘ 3.26x 25 (81.4)	↘ 0 (0)	↘ 17.46x 11 (192)	↘ 0 (44.3)	Baseball protective gear 491	↗ 2x 113 (56.6)	↘ 4.38x 3 (13.1)	↘ 4.32x 12 (51.8)	↘ 1.01x 22 (22.1)	↗ 1.13x 331 (294)	↗ 1.47x 89 (60.5)	↘ 0 (0)	↘ 20.41x 7 (143)	↘ 16.49x 2 (33)
LABEL / CAPTIONS_WORDS_GENDERED	boy 4,431	female 1,029	girl 4,057	male 1,732	man 22,997	men 4,741	trans 2	woman 11,188	women 2,583																																																																												
Baseball park 1,061	↗ 1.39x 170 (122)	↘ 4.73x 6 (28.4)	↘ 10.18x 11 (112)	↘ 1.04x 46 (47.8)	↗ 1.22x 777 (635)	↗ 1.57x 206 (131)	↘ 0 (0.1)	↘ 30.87x 10 (309)	↘ 17.82x 4 (71.3)																																																																												
Toddler 691	↗ 5.06x 403 (79.6)	↘ 6.16x 3 (18.5)	↗ 3.8x 277 (72.9)	↘ 4.45x 7 (31.1)	↘ 7.13x 58 (413)	↘ 85.2x 1 (85.2)	↘ 0 (0)	↘ 2.45x 82 (201)	↘ 11.61x 4 (46.4)																																																																												
Batting glove 639	↗ 1.64x 121 (73.6)	↘ 8.55x 2 (17.1)	↘ 13.48x 5 (67.4)	↗ 1.01x 29 (28.8)	↗ 1.3x 496 (382)	↗ 1.12x 88 (78.8)	↘ 0 (0)	↘ 26.56x 7 (186)	↘ 14.31x 3 (42.9)																																																																												
Kickflip 635	↗ 3.13x 229 (73.2)	↘ 16.99x 1 (17)	↘ 5.15x 13 (67)	↗ 2.31x 66 (28.6)	↗ 1.51x 573 (380)	↘ 3.01x 26 (78.3)	↘ 0 (0)	↘ 26.4x 7 (185)	↘ 42.66x 1 (42.7)																																																																												
Baseball 1,246	↗ 1.81x 260 (144)	↘ 5.56x 6 (33.3)	↘ 5.48x 24 (131)	↘ 1.08x 52 (56.1)	↗ 1.18x 876 (745)	↗ 1.48x 228 (154)	↘ 0 (0.1)	↘ 19.08x 19 (363)	↘ 9.3x 9 (83.7)																																																																												
Skateboard deck 660	↗ 2.73x 208 (76.1)	↘ 8.83x 2 (17.7)	↘ 6.33x 11 (69.6)	↗ 2.35x 70 (29.7)	↗ 1.51x 598 (395)	↘ 3.26x 25 (81.4)	↘ 0 (0)	↘ 17.46x 11 (192)	↘ 0 (44.3)																																																																												
Baseball protective gear 491	↗ 2x 113 (56.6)	↘ 4.38x 3 (13.1)	↘ 4.32x 12 (51.8)	↘ 1.01x 22 (22.1)	↗ 1.13x 331 (294)	↗ 1.47x 89 (60.5)	↘ 0 (0)	↘ 20.41x 7 (143)	↘ 16.49x 2 (33)																																																																												

✉ Подписаться ▾

авторизуйтесь

Оставьте первый комментарий!



B I U S $\frac{1}{3}$ = = " </> { } [+]



0 COMMENTS





Курс по нейронным сетям

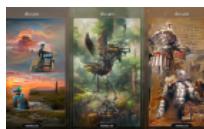
Используйте силу machine learning для решения задач бизнеса



Подписывайтесь:



Актуальное



Dream: приложение для создания стилизованных изображений

14 декабря 2021



Amazon представила инструменты AWS на основе машинного обучения

12 декабря 2021



RETRO: языковая модель DeepMind

10 декабря 2021



Intel анонсировала крупнейшие датасеты для распознавания речи

9 декабря 2021



PyTorch Live: фреймворк для создания мобильных приложений

6 декабря 2021

SkillFACTORY

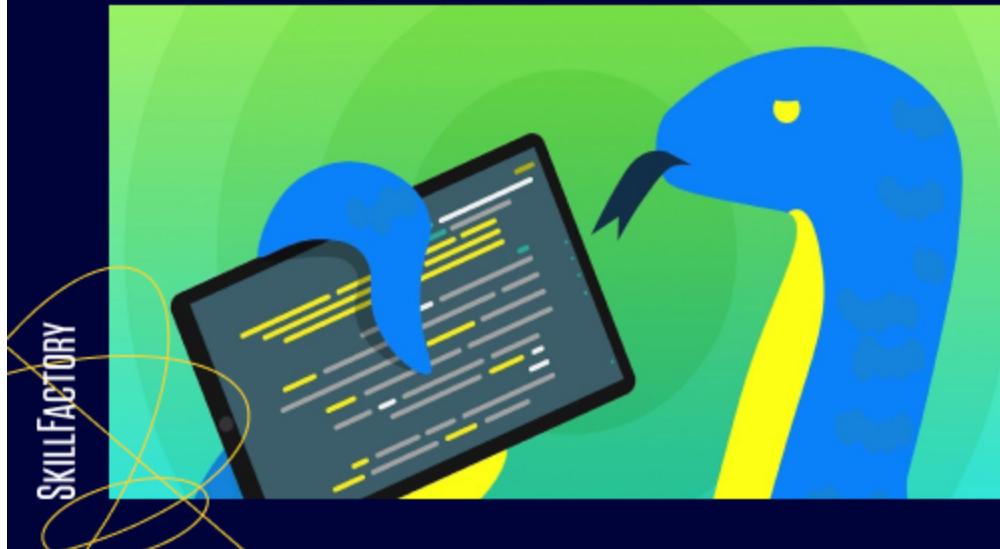
Математика и Machine Learning для Data Science

Сделайте свое резюме привлекательным для крупных компаний

курс

Научим использовать Python для анализа данных

Курс Python для анализа данных



Разрабатываем приложения и рассказываем о последних исследованиях в
области нейронных сетей: computer vision, nlp, обработка фотографий,
потокового видео и звука, дополненная и виртуальная реальность.

