

|  |
| --- |
| تعلم الآلة للبشر |
| فيشيل مانا |
| سامر صبري |
| ترجمة : محمود زرعي |

|  |  |  |
| --- | --- | --- |
| المصطلح | المعنى المستخدم |  |
| Machine learning | [تعلم الآلة](https://ar.wikipedia.org/wiki/%D8%AA%D8%B9%D9%84%D9%85_%D8%A7%D9%84%D8%A2%D9%84%D8%A9) |  |
| Supervised learning | [تعلم المشرف](https://ar.wikipedia.org/wiki/%D8%AA%D8%B9%D9%84%D9%85_%D9%85%D8%B1%D8%A7%D9%82%D8%A8) | تم تغير من مراقب إلى مشرف |
| Unsupervised learning | [تعلم بغير أشراف](https://ar.wikipedia.org/wiki/%D8%AA%D8%B9%D9%84%D9%85_%D8%BA%D9%8A%D8%B1_%D9%85%D8%B1%D8%A7%D9%82%D8%A8) |  |
| Reinforcement learning | [تعلم المقوى](https://ar.wikipedia.org/wiki/%D8%AA%D8%B9%D9%84%D9%8A%D9%85_%D9%85%D8%B9%D8%B2%D8%B2) | مقوى انسب من ناحيتي  حيث   [fortified] اقرب ل)معزز( من[ reinforcement] |
| \ cost function | أقتران تكلفة |  |
| Loss functions | اقتران الخسارة | اجتهاد شخصي – يرجى تعزيز او تغيرلمعنى اكثر مناسبة |
| Overfitting | إكداء | اجتهاد شخصي – يرجى تعزيز او تغيرلمعنى اكثر مناسبة |
| Underfitting | خمول | اجتهاد شخصي – يرجى تعزيز او تغيرلمعنى اكثر مناسبة |
| Regression | حسر | اجتهاد شخصي – يرجى تعزيز او تغيرلمعنى اكثر مناسبة |
| Classification | تصنيف |  |
| Gradient descent | هبوط تدرج | اجتهاد شخصي – يرجى تعزيز او تغيرلمعنى اكثر مناسبة |
| Logistic regression | الحسر لوغرثمي | Logistic عبارة عن [***log****istic un****it***] |
| Support vector machines (SVM) | تأيد متجهة الآلة \_ \_(امل) | اجتهاد شخصي – يرجى تعزيز او تغيرلمعنى اكثر مناسبة (يرجى الاحتفاظ بأمكانية انشاء تصغير) |
| Non-parametric learners | تعلم بدون معملات مدخلات |  |
| K-nearest neighbors (k) | الجار الأقرب لـ(ك) (جارك) | اجتهاد شخصي – يرجى تعزيز او تغيرلمعنى اكثر مناسبة (يرجى الاحتفاظ بأمكانية انشاء تصغير) |
| Decision trees | اشجار القرار |  |
| Random forests | الغابات العشوائية |  |
| Cross-validation | تصديق متقاطع |  |
| Hyperparameter tuning | ضبط المتغيرات الفائقة |  |
| Ensemble models | نموذج المجموعة |  |
| Agent | وكيل \ عميل \ **منظمومة** |  |
| Label | تسمية |  |
| [Tensor](https://en.wikipedia.org/wiki/Tensor) | [موتر](https://ar.wikipedia.org/wiki/%D9%85%D9%88%D8%AA%D8%B1) |  |
| Heuristics | الاستدلال |  |
| Ordinary least squares | طريقة القيمة الصغرى لمجموع المربعات |  |
| Sigmoid function | الأقتران سيني |  |
| Array | مصفوفة |  |
| List | قائمة |  |
| Tuples | مرصوصة |  |
| Set | مجموعة |  |
| Function | اقتران |  |
| Variance | تباين |  |
| bias | انحياز |  |
| threshold | عتبة |  |
| regularized cost | تكلفة تسوية |  |
| Markov decision processes | قرارات عملية ماركوف (قمم) | اجتهاد شخصي – يرجى تعزيز او تغيرلمعنى اكثر مناسبة (يرجى الاحتفاظ بأمكانية انشاء تصغير) |
| Non-parametric | بدون معملات |  |
| mode | منوال |  |
| Clustering | تحليل عنقودي (تجميع العنقودي) |  |
| k-means | خوارزمية تصنيفية |  |
| hierarchical clustering | تجميع هرمي |  |
| Dimensionality reduction | تخفيض الأبعاد |  |
| principal components analysis | تحليل العناصر الرئيسي (تار) | اجتهاد شخصي – يرجى تعزيز او تغيرلمعنى اكثر مناسبة (يرجى الاحتفاظ بأمكانية انشاء تصغير) |
| singular value decomposition | مجزئ القيمة المفردة (جيم) | اجتهاد شخصي – يرجى تعزيز او تغيرلمعنى اكثر مناسبة (يرجى الاحتفاظ بأمكانية انشاء تصغير) |
| Neural Networks | شبكات العصبية |  |
| Deep Learning | تعلم العميق |  |
| Convolutional neural networks | شبكات العصبية التفافية (شبت) | اجتهاد شخصي – يرجى تعزيز او تغيرلمعنى اكثر مناسبة (يرجى الاحتفاظ بأمكانية انشاء تصغير) |
| recurrent neural networks | شبكات العصبية المتكررة (شبر) | اجتهاد شخصي – يرجى تعزيز او تغيرلمعنى اكثر مناسبة (يرجى الاحتفاظ بأمكانية انشاء تصغير) |
| Q-learning | تعلم-كيو |  |
| Underfitting | ارخاء |  |
| feature | ميزة |  |
| Log-odd | الأرجحية – الوغرثمية |  |
| hyperplane | مسطح |  |
| Markov Decision Processes | قرارات عملية ماركوف (رما) | اجتهاد شخصي – يرجى تعزيز او تغيرلمعنى اكثر مناسبة (يرجى الاحتفاظ بأمكانية انشاء تصغير) |
| logit | لوغرثمي |  |
| sigmoid | الاقتران السيني |  |
| ثابت اويلر | **هــ** |  |
| بيتا | **ب** |  |
| ابسلون | **ي** |  |
| normalization | تسوية |  |
| Regularization | تنظيم |  |

الفهرس



الفصل 1: مقدمة. الصورة الكبيرة للذكاء الاصطناعي التعلم الآلة - الماضي والحاضر والمستقبل.

الفصل 2.1: تعلم المُشرف. التعلم بوجود الإجابات. إدخال الانحدار الخطي ، اقتران الخسارة ، الاكداء ، هبوط تدرج.

الفصل 2.2: التعلم المُشرف -2- طريقتان للتصنيف: الحسر الوغرثمي تأيد متجهة الآلة (امل).

الفصل 2.3: التعلم المُشرف-3-. تعلم بدون معملات مدخلات : الجار الأقرب لـ(ك) (جارك) ، أشجار القرار ، الغابات العشوائية. مقدمة في تصديق متقاطع ضبط المتغيرات الفائقة ونماذج المجموعات.

الفصل 3: التعلم غير الخاضع للإشراف تحليل عنقودي : خوارزمية تصنيفية ، تجميع هرمي .  
 تخفيض الأبعاد : تحليل العناصر الرئيسي (تار) , مجزئ القيمة المفردة (جيم).

الفصل 4 : الشبكات العصبية والتعلم العميق. لماذا وأين وكيف يعمل التعلم العميق. الإلهام من الدماغ. شبكات العصبية التفافية (شبت) ، شبكات العصبية المتكررة (شبر) , تطبيقات العالم الحقيقي.

الفصل 5 : تعليم المقوى . الأستكشاف و الأستغلال, قرارات عملية ماركوف , تعلم سياسة و تعلم العميق المقوى ومشكلة تعلم القيمة

الملحق: أفضل موارد التعلم الآلة. قائمة منسقة من الموارد لإنشاء منهج التعلم الآلة الخاص بك.

Machine Learning for Humans

الفصل الأول : المقدمة.



لمن هذا الكتاب ؟



الأشخاص التقنين الذين يريدون إن يطلوع على تعلم الآلة

الأشخاص الغير تقنين الذين يريدون إن يبنوى أساس معرفي عن تعلم الآلة وهم على استعداد على

الخوض في تفاصيل التقنية



أي شخص لديه فضول حول آلية تفكير الآلات

هذا دليل معد ليكون سهل المنال و الفهم لأي شخص . سيتم مناقشة مفاهيم مبسطة في الآحتمالات , الأحصاء , البرمجة , الجبر الخطي و تفاضل وتكامل ,ولكن ليس هناك أي ضرورى لأمتلاك معرفة مسبقة حول هذه المواضيع , لأخذ الفائدة من هذا الكتاب .

إذا كنت راغب بمعرفة أي دورات تدريبة , الكتب و المشاريع قم بألقاء نظرة على الملحقات لمعرفة أختياراتنا حول أفضل المصدار لتعلم حول تعلم الآلة

لماذا تعلم الآلة يهّم ؟

سيساهم الذكاء الأصطناعي بتشكيل مستقبل البشرية بدرجة لم يسبق لأي اختراع إن فعلها , سنستيقظ في عالم مليىء بتقنيات أقرب إلى الخيال .

معدل تسارع في هذا المجال مثيرة للأعجاب , وخصوصاً بعد عدد من فترات [سبات و الأمل زائف](https://en.wikipedia.org/wiki/History_of_artificial_intelligence" \l "The_first_AI_winter_1974.E2.80.931980) خلال الأربع عقود المنصرمة , وتطور المتسارع في سعات تخزينية و القدرات المعالجات التي ساعدت في تسريع نمو هذا المجال.

Machine Learning for Humans

في 2015 قامت جوجل بتدريب منظومة للمحادثة للقيام ليس فقط بتقديم دعم الفني ولكن مناقشة أمور مثل الأخلاق , حرية تعبيرو للأجابة عن الأسئلة العامة المبنية على الحقائق .



([Vinyals & Le, 2017](https://arxiv.org/abs/1506.05869))

Machine Learning for Humans

وفي نفس العام , قامت (DeepMind) بتطوير منظومة , وهذه [المنظومة](https://storage.googleapis.com/deepmind-media/dqn/DQNNaturePaper.pdf) قد تخطت المستوى البشري في اللعاب (49 Atari), رغم تلقيها بيكسيلات شاشة ونتيجة اللعبة فقط. وبعدها بفترة قصير قامت (DeepMind) بتفوق على نفسها عبر أطلاق [A3C](https://arxiv.org/pdf/1602.01783.pdf)

وفي نفس الوقت , قامت (آلفا غو) بهزيمة أفضل لاعب بشري في للعبة (غو) , انجاز فائق للعادة خصوصاً إن هذه اللعبة يسيطر فيها البشر ولآن بعد عقدين من فوز ذكاء الأصطناعي في للعبة شطرنج. الكثير من محترفي اللعبة لا يستطيعون فهم كيف لآلة إن تدرك كامل تعقيدات وخفاية هذه اللعبة الأستراتيجية الصينية العتقية , ونظراً لآنه يوجد 10170 وضعية للأحجار على رقعة وضعاً بالعلم إنه يوجد [10](http://www.slate.com/articles/technology/technology/2016/03/google_s_alphago_defeated_go_champion_lee_sedol_ken_jennings_explains_what.html)[80](http://www.slate.com/articles/technology/technology/2016/03/google_s_alphago_defeated_go_champion_lee_sedol_ken_jennings_explains_what.html) [ذرة بالكون المرئي](http://www.slate.com/articles/technology/technology/2016/03/google_s_alphago_defeated_go_champion_lee_sedol_ken_jennings_explains_what.html).



بطل العالم في لعبة (غو) يراجع حركاته بعد هزيمته على يد الفاغو

Machine Learning for Humans

وفي مارس 2017 , قامت (OpenAi) بصنع منظومات قامت [ببتكار لغتها الخاصة](https://blog.openai.com/learning-to-communicate/) لتتعاون و تحقق أهدافها بشكل أكثر فعالية. وليس بمدة طويلة قامت (FaceBook) بصنع منظومة تمكنت من [المفاوضة](https://code.facebook.com/posts/1686672014972296/deal-or-no-deal-training-ai-bots-to-negotiate/) وحتى [الكذب](https://www.theregister.co.uk/2017/06/15/facebook_to_teach_chatbots_negotiation/).

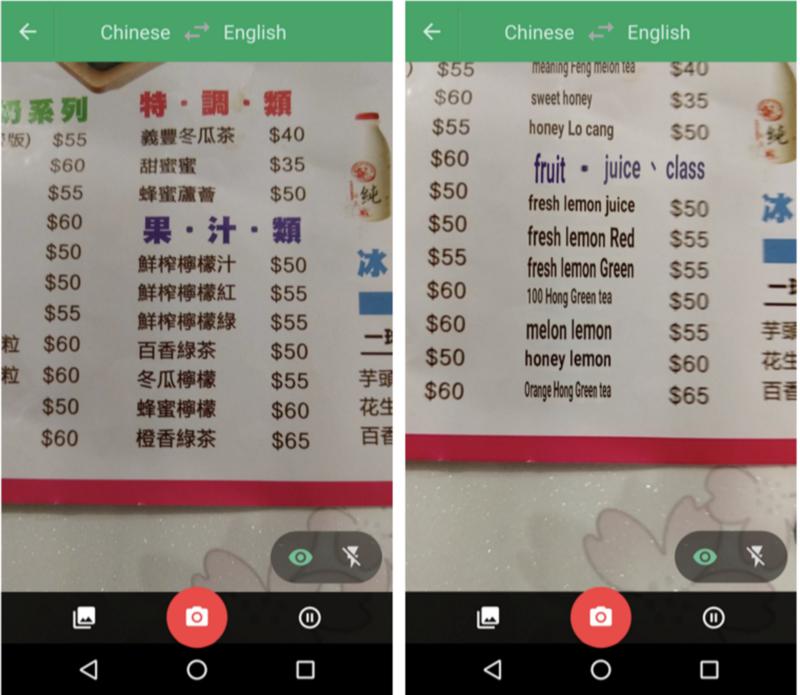
وفي الحادي عشر من أغسطوس 2017 , قامت (OpenAi) بتحقيق انجاز أخر, عبر هزيمة اللاعبين المصنفين في المرتبة الأولى في للعبة دوتا 2



Machine Learning for Humans

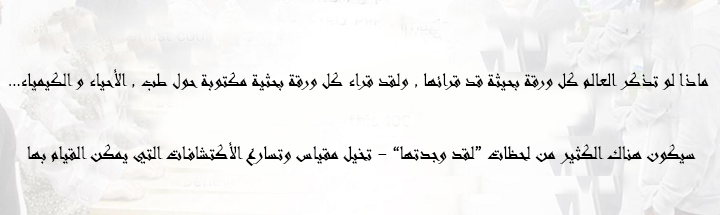
الكثير من تقنيات التي تُسير حياتنا اليومية تعتمد على الذكاء الأصطناعي , قم بلتقاط صورة للائحة طعام باللغة صينية

وبشكل سحري تظهر باللغة الأنجليزية بواسطة تطبيق ترجمة الخاصة بقوقل .



Machine Learning for Humans

يتم استخدام الذكاء الاصطناعى اليوم لتصميم [خطط العلاج القائمة على الأدلة لمرضى السرطان](https://www.ibm.com/watson/health/oncology-and-genomics/oncology/) ، وتحليل النتائج على الفور من الاختبارات الطبية [وتوكيلها إلى الاختصاصي](https://deepmind.com/applied/deepmind-health/) المناسب على الفور ، [وإجراء البحوث العلمية](http://benevolent.ai/) لاكتشاف الأدوية.



تستخدم قوات شرطة , تقنية تعرف على الأوجه و معالجة اللغة الطبيعية [لمعاينة صور الملتقطة من خلال كاميرات الجسم](https://www.axon.com/company/news/axon-ai-work).

وكما يستخدم جوال المريخ (Curiosity) ذكاء الأصطناعي لأتمتة عملية البحث عن [عينات جديرة تستحق الجمع والبحث](http://www.astronomy.com/news/2016/08/how-does-mars-rover-curiositys-new-ai-system-work) بدقة عالية

وفي الحياة اليومية , يتم بشكل متزايد استخدام ذكاء الأصطناعي في العديد من الأمكان الشائعة والوظائف التي يشغلها البشر . لا تكن متفاجىء عندما يقرع بابك روبوت توصيل عوضاً عن رجل توصيل .

في هذا دليل , سنقوم بشرح المفاهيم الأساسية تعلم الآلة , في نهاية . يجب إن تكون قادر على وصف بشكل فكري كيفية عمل تعلم الآلة وستكون قادر على البدء بصنع تطبيقات مشابه بنفسك.

الشجرة الدلالية: الذكاء الاصطناعي والتعلم الآلي

*نصيحة واحدة: من المهم النظر إلى المعرفة على أنها نوع من الشجرة الدلالية*

*- تأكد من فهمك للمبادئ الأساسية ، أي الجذع والأغصان الكبيرة*

*، قبل الدخول في الأوراق / التفاصيل أو لا يوجد شيء يمكنهم تعليق به*

*اليون ماسك : ريديت أسالني أي شيء*

Machine Learning for Humans



تعلم الألة هو عبار عن فرع من فروع ذكاء الأصطناعي , يهتم بكيفية تعلم الحواسيب من الخبرة لتحسين قدراته على تفكير , تخطيط , اخذ القرار وتصرف

**ذكاء الأصطناعي : عملية دراسة المنظومات التي تستشعر العالم من حولها , من مستوى , وتقوم بصنع القرارات لتحقيق أهدافها.**

اساسات هذا العلم تتضمن : الرياضيات , المنطق , الفلسفة , الأحتمالات , اللغويات , علم الأعصاب , نظرية القرارات . والمزيد من الفروع التي يشملها ذكاء الأصطناعي مثل رؤية الحاسوب , الروبوتات , تعلم الآلة , ومعالجة اللغة الطبيعية.

**تعلم الآلة هو فرع من ذكاء الأصطناعي. وهدفه هو تمكين الحواسيب من تعلم بمفردها.** خورزميات تعلم الآلة تمكن الحواسيب من تميز الأنماط المخفية في البيانات , بناء نماذج تقوم بشرح العالم , وكما تمكنها من توقع الأحداث بدون برمجتها بشكل مسبق.

Machine Learning for Humans

تأثير الـ(ذكاء الأصطناعي) : ما شيء الذي يحق إن يطلق عليه ذكاء أصطناعي

تعريف دقيق لتقنية تأهلهت لتصبح ذكاء صناعي تعريف مشوش قليلاً. وتفسيرهات تغيرت خلال الوقت . مسمى ذكاء الأصطناعي يميل لوصف الآلات التي تقوم بأنجاز مهمات التي يستطيع البشر القيام , وعندما يستطيع الذكاء الأصطناعي القيام بهذه المهمة بأداء أفضل أو موازي لأداء البشر , يعترض البشر ويقولون هذا ليس "ذكاء" وهذا ما يعرف [بتأثير ذكاء الأصطناعي](https://en.wikipedia.org/wiki/AI_effect)

وعلى سبيل المثال , عندما قام (Deep Blue) الخاص بـ(IBM) بهزيمة بطل العالم في شطرنج [جاري كاسبروف](https://medium.com/u/66cdbe693a4f) في 1997

اشتكى ناس إن هذا الفعل تم بستخدام طريقة "القوى الفارطة" ولذالك لا يتأهل لأن يكون ذكاء أصطناعي "حقيقي". كما كتبت باميل مككوردوك." أنها جزء لا يتجزء من تاريخ مجال الذكاء الأصطناعي , في كل مرة يقوم شخص بكتشاف طريقة لجعل الحسوب يقوم بشيء ما ,مثل القيام بلعب أكس أو أو بحل مسال بسيطة سيكون هناك جمهور من نقاد يغنون "[هذا ليس ذكاء أو تفكير](http://www.pamelamc.com/html/machines_who_think.html)"

ربما يوجد شيء ما *لا أعرف مثل ماذا* سيقوم ناس بعتباره ذكاء أصطناعي :

ذكاء الأصطناعي هو شيء الذي لم يتم تحقيقه للأن – دوجلوس هوفستادتر

هل هذا يعني إن الآلة الحاسبة تعتبر ذكاء أصطناعي؟ ربما حسب بعض تفسيرات . ماذا عن سيارة ذاتية القايدة ؟ اليوم نعم انها تعتبر ذكاء اصطناعي , ولكن في المستقبل قد لا تعتبر. وماذا عن برنامج المحادثة الخاص بك . نعم لما لا.

ذكاء الأصطناعي العام سيغير عالمنا للأبد : لفهم كيف , البدء في دراسة تعلم الآلة نقطة مناسبة للبدء

تقنيات التي تم مناقشتها سابقاً تقع تحت خانة الذكاء الأصطناعي المحدود , والذي بدوره يستطيع القيام بمهمات محدودة ومحددة

Machine Learning for Humans

وفي هذه الأثناء, العالم يقوم بمواصلة الرحلة نحوى ذكاء اصطناعي بمستوى بشري , أو كما يطلق [عليه ذكاء الأصطناعي العام](https://en.wikipedia.org/wiki/Artificial_general_intelligence) . يعرف ذكاء الأصطناعي العام بأنه ذكاء الأصطناعي الذي يستطيع تأدية المهمات التي يستطيع البشر القيام بها بنجاح , ومنها تعلم ,تخطيط , اتخاذ القرارت تحت اي ظرف , تواصل بستخدام اللغة طبيعية , القاء نكات , تلاعب بالبشر , تجارة , اعادة برمجة نفسه.

وأعادة برمجة هي أكثرهن أهمية , إذا استطعنا القيام بأنشاء ذكاء اصطناعي يستطيع تحسين من نفسه بستمرار , سيقوم هذا بخلق حلقة من تحسين المستمر مما يؤدي إلى **انفجار ذكاء** خلال مدة غير معلومة من الوقت قد تتراوح من عدد من العقود إلى يوم واحد.

*دع آلة فائق للذكاء يمكن تعريفها بأنها آلة يمكنها أن تفوق في كل الأنشطة الفكرية على الانسان مهما كان ذكيا. بما أن تصميم الآلات هو واحد من هذه الأنشطة الفكرية ، يمكن للآلة فائقة الذكاء تصميم آلات أفضل. عندئذ سيكون هناك بلا شك "انفجار في ذكاء" ، وسيتم ترك ذكاء الانسان للوراء. وبالتالي ، فإن أول آلة فائقة الذكاء هي الاختراع الأخير الذي يحتاج إليه الإنسان ، شريطة أن يكون الجهاز قابلاً بما فيه الكفاية ليخبرنا كيف نحتفظ به تحت السيطرة. - اي جاي. قود ، 1965*

ربما قد سمعت عن مصطلح يسمى **المنفردة**. هذا المصطلح تم أستعارته من الفيزياء ويعبر هذا المصطلح عن انفراد الجاذبية الذي يحدث في مركز ثقب الأسود ,نقطة احادية الأباعد لها كثافة لا نهائية , عندها تتوقف الفيزياء عن العمل.

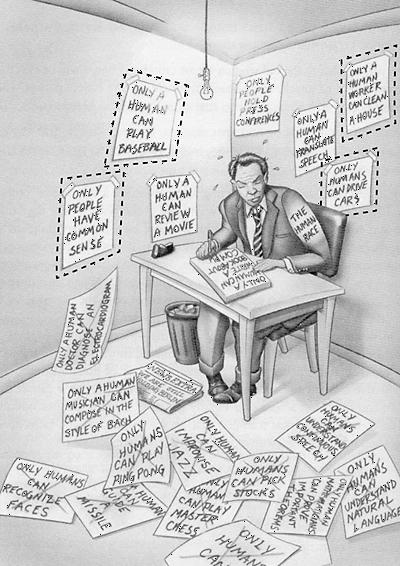
Machine Learning for Humans



وليس لدينا ادنى فكرة عما يحدث بعد آفق الحدث الخاص بثقب الأسود بسبب إن ضوء لا يستطيع الخروج منه. وبشكل مشابه ,عندما نعطي الذكاء الأصطناعي القدرة تحسين ذاته بشكل مستمر , سيكون من المستحيل تنبىء بما سيحدث حينها , مثل من اخرج الجن من قمقمه لا احد يعرف ماذا سيفعل الجن. هل سيحقق أمنياتك ؟

قامت مؤسسة "مستقبل البشرية" بنشر تقرير قامت به بستطلاع أرىء الباحثين في الذكاء الأصطناعي حول المدى الزمنية حتى نصل لمرحلة ذكاء الأصطناعي العام . "[حيث كشف تقرير إن الباحثين يأملون إن هناك فرصة 50% إن ذكاء الأصطناعي سيقوم بتأدية كل المهمات القادر على أدائها البشر خلال 45 عام](https://arxiv.org/pdf/1705.08807.pdf) ". ونحن بشكل شخصي قد حاورنا عدد من المتمرسين في هذا المجال , والذين بدورهم يتوقعون وقت أطول من 45 عام وبعضهم يقول إنه من المستحيل حدوثه بالأساس والبعض يقول إن الوقت اقل من بضع سنين.

Machine Learning for Humans



صورة من كتاب رآي كيرتزول المنفردة قريبىة . الذي نشر في 2005 وفي وقتنا فقط القليل من الملصقات ستبقى على الحائط

اختراع ذكاء الأصطناعي الفائق , الذي سيكون اذكى من مجموع ذكاء الجنس البشري . قد يكون من افضل الأمور التي حدثت للجنس البشري , ولكنه امر صعب القيام ببتكار هذا ذكاء وجعله ودي اتجاه البشر.

ومعرفة ماذا يخفي المستقبل هو أمر مستحيل , يوجد شيء مؤكد إن عام 2017 هو عام رائع لبدء فهم كيف تفكر الآلات.

لتتعمق بعد تجرود الفلسفة و لنقم بذكاء بتشكيل طريقنا بأخذ الأعتبار لذكاء الأصطناعي , يجب إن نشارك تفاصيل كيفية رؤية الألات للعالم , لنرى ما يريدون ولنعرف تحيزات التي تؤدي لفشل نموذج , الميزات تلاقائية كما ندرس علم نفس وعلم الأعصاب لفهم كيفية إن بقوم البشر بتعلم , اتخاذ القرارات , وشعور.

Machine Learning for Humans

هناك أسئلة معقدة ومحفوفة بالمخاطر حول الذكاء الاصطناعي ستتطلب اهتمامنا الدقيق في السنوات القادمة.

كيف يمكننا مكافحة نزوع الذكاء الاصطناعي إلى ترسيخ التحيزات المنهجية الواضحة في مجموعات البيانات الحالية؟ ما الذي يجب أن نستخلصه من الخلافات الأساسية بين أقوى خبراء التكنولوجيا في العالم حول المخاطر والفوائد المحتملة للذكاء الاصطناعي؟ ما هي الأساليب التقنية الواعدة لتعليم أنظمة الذكاء الاصطناعي كيف تتصرف؟ ماذا سيحدث لإحساس البشر بالهدف في عالم بلا عمل؟

تعلم الآلة يعتبر حجر الأساس لبدء طريق في تعلم عن الذكاء الأصطناعي , وفي هذه الأثناء يقوم الذكاء الأصطناعي بتغير كل صناعة وحرفة وسيكون له تأثر على حياتنا اليومية بشكل لا يمكن تخيله , ولهذا نعتقد إنه من المهم إن تفهم كيف يعمل تعلم الآلة , على الأقل على المستوى الفكري , وهذا دليل افضل مكان للبدء.

كيف تقراء هذا الكتيب

ليس من ضروري إن تقراء هذا دليل بشكل كامل حتى تستخلص الفائدة منه . يوجد ثلاث طرق مقترحة وكل طرق تعتمد على رغباتك و الوقت المتاح إليك:

1. **مقاربة شكل ( (T** : القرائة من البداية إلى نهاية . قم بتلخيص كل فصل بستخدام افكارك وكلماتك (شاهد طريقة [فايمن](https://mattyford.com/blog/2014/1/23/the-feynman-technique-model)), هذا يشجع القرائة الفعالة ويساعد على الاحتفاظ بالمعلومات لوقت اطول. تعمق في المجالات الأكثر صلة بعملك. سنقوم بتضمين موارد في نهاية كل قسم.
2. **المقاربة المركزة:** اقفز مباشرة إلى الفصول التي تثير فضولك وركز طاقتك الذهنية هناك.

**3. مقاربة 20\80** : قم بالقاء نظرة سريعة , دون الأشياء التي جذبت انتباهك.

Machine Learning for Humans

الفصل ثاني الباب الأول : تعليم المشرف



الوظيفتان التي يقوم بهما تعليم المشرف : التصنيف و الحسر. الحسر الخطي , آقتران الخسارة و هبوط تدرج.

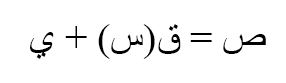
كم عائد الأرباح التي سنحصل عليها إذا انفقنا المزيد على الأعلانات الرقمية؟هل سيدفع المُقترض القرض أم لا؟ماذا سيحصل في سوق الأسهم غداً؟

في تعليم المشرف , نقوم بالبدء بستخدام مجموعة من البيانات تحتوي على أمثلة لتدريب مع تسميات صحيحة.على سبيل المثال,عند تعلم كيفية تصنيف الأرقام المكتوبة بخط اليد, خورزمية تعليم المشرف تقوم بأخذ آلاف من صور الأرقام المكتوبة بخط اليد بحيث يكون اسم صورة هو رقم الموجود في صورة .

الخورزمية ستقوم بتعلم محتوى صورة وربطه بأسم صورة , وتقوم بتطبيق العلاقة التي تعلمتها على صور جديدة (مع اسماء) , وهكذا يتم إيداع شيك عبر التقاط صورة له.

لتوضيح كيف يعمل تعلم المشرف , سنفحص معضلة توقع دخل السنوي بنائاً على عدد سنين تعليم العالي التي تلقها شخص وبصيغة أخرى سنقوم ببناء نموذج يقوم بتقدير العلاقة بين عدد سنين تعليم العالي و دخل السنوي .

Machine Learning for Humans



**س (المدخلات) : عدد سينين تعليم العالي**

**ص (المخرجات) : دخل سنوي**

**ق : آقتران يصف العلاقة بين س و ص**

**ي (إبسليون) : قيمة عشوائية تقوم بتقدير كل شيء لا يقوم نموذج بوضع الحسبان له**

بخصوص ابسلون

1. ابسلون تمثل **الخطاء الذي لا يمكن تكراره** في نموذج, والذي يمثل سقف اداء نموذج نظراً لوجود ضوضاء في ظاهرة التي تحاول تفسيرها , على سبيل المثال , تخيل بناء نموذج يقوم بتوقع نتيجة رمي قطعة نقدية

وبشكل متعمد كان الرياضي بول ايردوس يشير إلى الأطفال بـ(ابسلون) لانه في تفاضل وتكامل (وليس في الأحصاء) أبسلون تشير أى قيمة عشوائية صغير موجبة , ملائم ؟ اليس كذالك . واحدة من طرق توقع دخل , ستكون عبر طرق نماذج مبنية على قواعد شامخة لوصف كيف يتعلق دخل بسنوات تعلم , على سيبل المثال : " انا اتوقع بانه لكل سنة من سنوات تعليم العالي , دخل سنوي يزداد بـمقدار 5000$)

دخل = ($5000\*عدد سنين تعليم العالي) + دخل الأساسي

هذه المقاربة مثال على **هندسة** الحل (على نظير **تعلم** الحل بمثل طريقة الحسر الخطي التي سنتناولها تالياً)

نموذج اكثر تعقيداً من خلال الأخذ بالحسبان بعض القواعد حول نوع , عدد سنوات الخبرة العملية , مرتبة الجامعة . على سبيل المثال:"إذا كان يمتلك شهادة بكالوريوس أو أعلى , قم بمضاعفة راتب المتوقع بمقدار 1.5"

ولكن هذا نوع من برمجة القواعد الواضحة وصريحة لا يترجم بشكل جيد عند تعامل مع البيانات المعقدة.تخيل إنك تقوم بتصميم خوارزيمة تقوم بتصنيف صور بستخدام الجمل شرطية , واصفًا كل تركيبة من البيكسلات بختلاف درجات سطوع لتصنيف صورة إلى "قط" أم "لا".

Machine Learning for Humans

تعلم المشرف يقوم بحل هذه المشكلة من خلال جعل الحاسوب يحل هذه المسألة بنيابة عنك. بواسطة تعرف على الأنماط في البيانات, الآلة قادرة على تشكيل حدس مهني.الأختلاف رئيسي بين هذا و تعلم الآنسان هو إن تعلم الآلة يتم بواسطة عتاد جهاز الحاسوب , ومن المفضل إن ننظر لهذا الأختلاف من منظور علم الحاسوب و الأحصاء, حيث يقوم الأنسان بتوصيل الأنماط بواسطة دماغ بيولويجي (ولأثنان (العتاد و دماغ) يصلاً على نتيجة متقاربة)

في تعليم الآلة المشرف , تحاول الآلة إن تتعلم العلاقات بين دخل ودرجة تعليم عالي من صفر . بواسطة تمرير **البيانات المسماة** من خلال **خوارزيمة تعلم** . وهذا العلاقة (الأقتران) تستخدم لتقدير دخل ص الأشخاص في حال تواجد عدد سنين تعليم العالي س , بصياغة اخرى نقوم بتطبيق نموذجنا على **بيانات غير مسماة** لتقدير قيمة ص.

هدف تعليم المشرف هو **توقع قيمة ص بدرجة عالية من دقة على قدر المستطاع** . في تآلي سنقوم بستكشاف عدد من طرق لمقاربة هذه طريقة.

الوظيفتان التي يقوم بهما تعليم المشرف : تصنيف و الحسر

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| الحسر: | تصنيف: |  | تصنيف: |  |
| توقع قيمة عددية متصلة. | تعين تسمية. |  | تعين تسمية. |  |
| كم سيكون سعر هذا المنزل عند بيعه؟ | هل هذا قط أم كلب؟ |  | هل هذا قط أم كلب؟ |  |

في ما بقي من هذا الجزء سنقوم بتركيز على الحسر . في الجزء ثاني الباب ثاني سنقوم بتغول بعمق في طرق التصنيف.

الحسر: توقع قيمة متصلة

الحسر يقوم بتوقع قيمة عددية متصلة للمتغير المستهدف ص. هو يسمح لك بتقدير القيمة , مثل سعر المنازل , فترة حياة الأنسان اعتماداً على البيانات المدخلة س.

هنا,**المتغير المستهدف** هو المتغير صاحب القيمة المجهولة التي نحاول توقعها, و **المتصل** يعني انه لا يوجد فراغات في القيمة التي يمكن تعينها لـ(ص).

Machine Learning for Humans

وزن و طول الشخص يعتبران مثالين على القيم المتصلة , اما من ناحية اخرى توجد القيم **المتقطعة** التي تؤخذ قيمة محدودة من الأعداد مثل عدد الأطفال الذي يمكن لشخص إن يحظى بهم.

توقع دخل هو مثال على عميلة الانحدار . **لديك مدخلاتك (س)** التي تحتوي كل البيانات التي لها صلة عن الأشخاص في مجموعة البيانات , التي ستستخدم في توقع دخولهم مثل , عدد سنين تعليم , عدد سنوات الخبرة , المنصب , المنطقة جغرافية . هذه صفات تدى **خصائص** والتي قد تكون عددية مثل سنين الخبرة أو وصفية مثل المنصب او حقل دراسة.

تحتاج لكل بيانات تدريب المتعلقة بهذه صفات الخاصة بشخص الذي نريد توقع دخله , من أجل إن يتعلم نموذجك الأقتران بين س و ص

نقوم بفصل البيانات إلى بيانات تدريب وبيانات اختبار . بيانات تدريب تحتوي على مسميات حتى يتكمن نموذجك من تعلم من خلال هذه البيانات المسمة . اما بيانات الأختبار فلا يجب إن تحتوي على مسميات

**الحسر**

**ص = ق(س) + ي ,** س =(س1 , س2 , .... , سن)

**تدريب :** الآلة تقوم بتعلم ق من البيانات المسماة

**الأختبار :** الآلة تقوم بتوقع ص من البيانات الغير مسمى

الاحظ إن س قد تكون موتر باكثر من بعد , موتر ببعد واحد يسمى متجه (صف واحد و أكثر من عمود) , موتر ببعدين يسمى مصفوفة (العديد من صفوف والعديد من الأعمدة) ويمكن للموتر إن يمتلك 3 , 4 و خمس ابعاد او اكثر . لمراجعة هذه المصطلحات

[مراجعة الجبر الخطي](http://www.deeplearningbook.org/contents/linear_algebra.html).

في مثالنا ثنائي الأبعاد المبسط , قد تكون البيانات على شكل بيانات مفصولة بفاصلة (CSV)حيث يحتوي كل صف على تعليم شخص و دخل .يمكنك اضافة المزيد من الأعمدة مع المزيد من **الخصائص** للحصول على نمذج اعقد ولكن ادق.

Machine Learning for Humans



إذا كيف سنحل هذه المشاكل ؟

**كيف سنبني نموذج , يستطيع القيام بتوقعات دقيقة و مفيدة , في الحياة الواقعية ؟ سنقوم بعل ذالك بستخدام خورزمية تعليم المشرف**

**الأن لنبدء بالجزء الممتع : تعرف على الخورزميات** , سنتعرف على طرق لمقاربة الحسر و تصنيف . وسنقوم بتوضيح مبادء اساسية في تعلم الآلة خلال ذالك

Machine Learning for Humans

الحسر الخطي (طريقة القيمة الصغرى لمجموع المربعات)

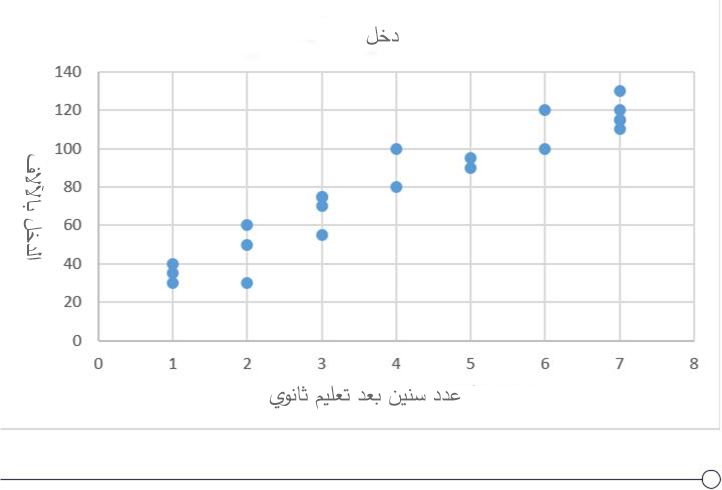
*“رسم خط, نعم هذا يعتبر تعلم الآلة.”*

في بادء الأمر , يجب إن نركز على حل مشكلة توقع دخل بستخدام الحسر الخطي , لان نماذج الخطية لا تعمل بشكل جيد مع المهام المتعلقة بتعرف على صور . (تعرف على صور هو مجال تعلم العميق , الذي سنستكشفه بشكل لاحق)

**لدينا مجموعة بيانات س , التي تتلائم مع قيمة ص , هدف** طريقة القيمة الصغرى لمجموع المربعات هو تعلم نموذج الخطي , حيث يمكن توقع قيمة ص إذ اعطي قيمة س جديدة . بأقل نسبة خطأ ممكنة . نريد إن نتوقع مقدار دخل الذي يجنيه شخص بناءا على عدد سنوات تعليم

س تدريب = [4 , 5 ,0 ,2 , ... , 6 ] # سنوات تعليم بعد تعليم ثانوي

ص تدريب = [80 , 91.5 , 42 , 55 , ... ,100]# دخل سنوي , كل وحددة توازي الف دولار



Machine Learning for Humans

الحسر الخطي هو طريقة بستخدام معاملات , وهذا يعني انها تقوم بصنع تخمينات حول شكل القتران الذي يربط س وص

(بالعربية (**parametric method**) تتطلب معاملات مثل العمر و دخل (non-parametric) لا تتطلب معملات )



في هذه الحالة , نحن نقوم بتخمين صريح إنه يوجد **علاقة خطية** بين س و ص , وانه بكل زيادة في س يوجد زيادة في ص (أو نقصان)

ب0 هي تقاطع مع محور ص , ب1 هي ميل الخط , الذي يعبر عن زيادة او نقصان دخل لكل سنة من سنوات تعليم

هدفنا هو تعلم معطيات نموذج التي تقليل من قيمة خطـأ هذا نموذج.

لأيجاد افضل معطيات:



1. قم بتعريف اقتران تكلفة\* او الخسارة\* , حيث يقوم هذا الاقتران بقياس مدى دقة توقع نموذجك



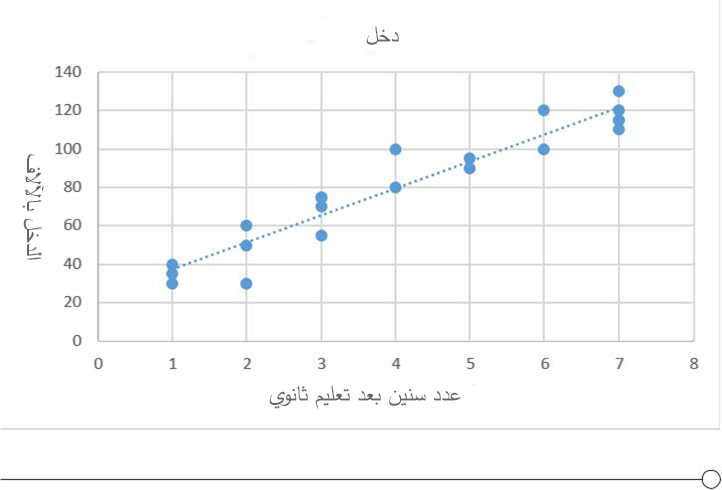
1. قم بالبحث عن المعاملات التي تقلل قيمة اقتران تكلفة , لنجعل توقعات نموذج دقيقة بأكبر شكل ممكن

قام الكاتب بفتراض بان اقتران تكلفة و اقتران الخسارة مكافئين ولكن في الوقع اقتران تكلفة هو عبار عن مجموع اقترانات الخسارة لجميع بيانات تدريب

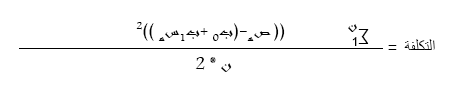
صورياً , في بعدين , هذا ينتج خط يلائمه . في ثلاثة ابعاد ينتج مسطح , وفي الأبعاد العلية ينتج مستو فائق.

*.* *ملاحظة عن البعدية: مثالنا ثنائي الأبعاد للبساطة ، ولكن عادةً ما يكون لديك المزيد من الميزات (س) ومعاملات (بيتا) في النموذج الخاص بك ، على سبيل المثال عند إضافة المزيد من المتغيرات ذات الصلة لتحسين دقة تنبؤات النموذج الخاص بك. تعمم نفس المبادئ على أبعاد أعلى ، على الرغم من صعوبة رؤية الأشياء بعد ثلاثة أبعاد.*

Machine Learning for Humans



رياضياً : نقوم بأخذ الفرق بين كل من نقاط الفعلية (ص) و توقعات نموذج (صْ) , نقوم بتربيع الفرق لتفادي الأرقام سالبة , ونقوم بجمعهم و أخذ المعدل . وهذا المعدل يقيس مدى مطابقة توقعهنا للخط



ن عدد نقاط تدريب . إن استخدام ن \* 2 بدلاً من ن يجعل الرياضيات اكثر اناقة عند أخذ المشتق لتقليل الخسارة إلى الحد الأدنى ، على الرغم من أن بعض الأحصائين يقولون بإن هذا نفاق. عندما تبدأ في إبداء الرأي حول هذا النوع من الأشياء ، ستعرف أنك على طول الطريق في حفرة الأرانب.

لحل مشكلة بسيطة مثل هذه , نقوم بحساب افضل قيمة ملائمة لبيتا التي تجعل اقتران تكلفة في اقل قيمة له بستخدام تفاضل وتكامل عبر حساب المساحة المغلقة لشكل , ولكن إذ كان اقتران تكلفة معقد فأن ايجاد القيمة الملائمة لبيتا بستخدام تفاضل وتكامل يصبح غير متاح وهذا كان الحافز وراء مقاربة تكرارية تسمى (**هبوط تدرج**) والتي تجعلنا نجد اقل قيمة لأقتران تكلفة معقد

Machine Learning for Humans

هبوط تدرج : تعلم المتغيرات

*"قم بوضع عصبة على عيناك , قم بأخذ خطوى نحو المنحدر , لقد وجدت القاع عندما لا يوجد مكان لذهاب سوى الأعلى"*

هبوط تدرج سيظهر في الكثير من المواضيع و بالأخص شبكات العصبية , و مكتبات تعلم الآلة مثل [scikit-learn](http://scikit-learn.org/stable/) و [TensorFlow](https://www.tensorflow.org/) تستخدم هبوط تدرج في كل مكان , لذالك يستحق فهم تفاصيل .

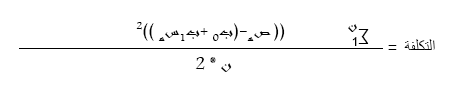
هدف هبوط تدرج هو ايجاد اقل قيمة لقتران تكلفة عبر محاولة تحسين تقريب لقتران تكلفة

تخيل نفسك تمشي معصوب العيني في واد , وهدفك هو ايجاد قاع هذا الواد , كيف ستفعل ذالك؟

طريقة منطقية ستكون إن تتحس الأرض حولك وتتوجه نحو الأتجاه الذي تنحدر فيه الأرض بشدة . خذ خطوى واعد الفعل حتى تصل إلى ارض مسطحة . وعندها تعرف انك وصلت قاع الواد : إذ تحركت في اي اتجاه ستنتهي في نفس الأرتفاع او اعلى قليلاً

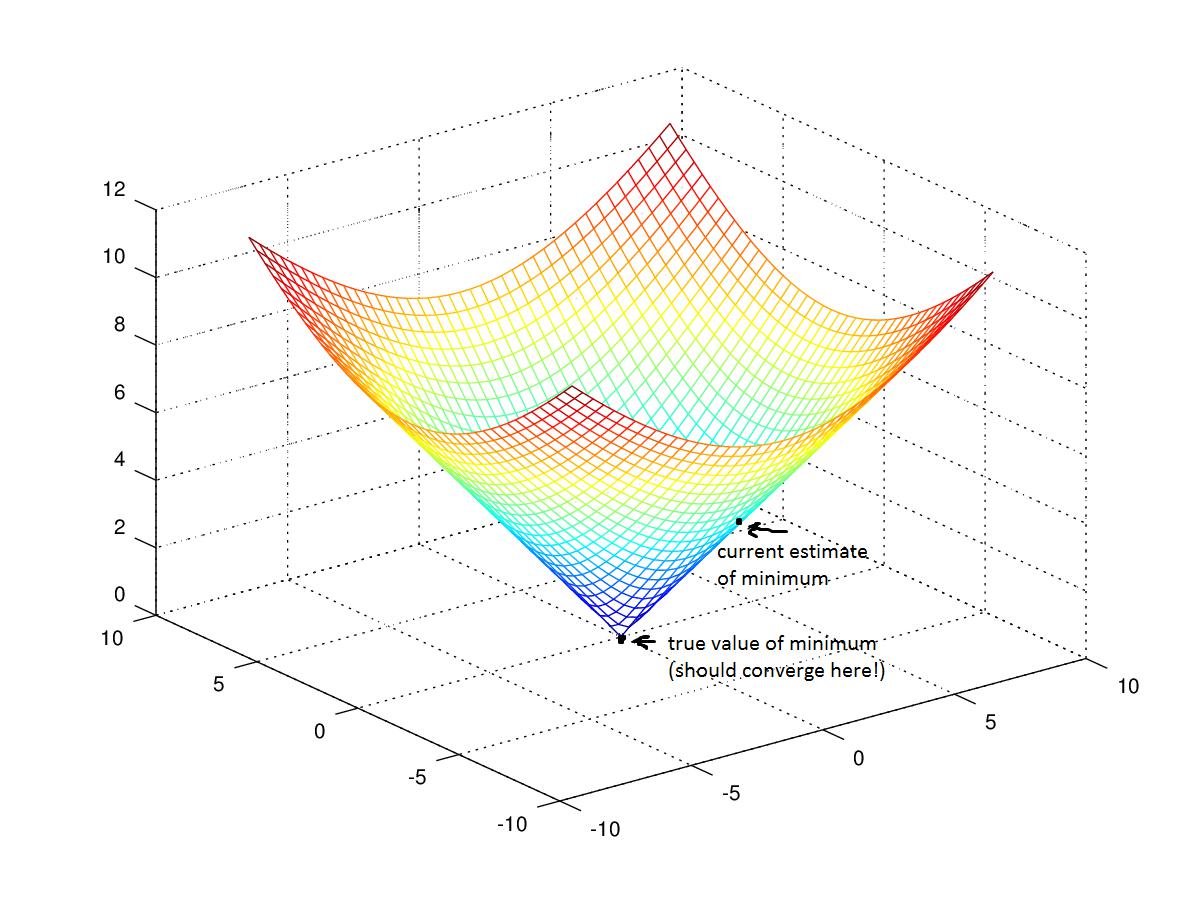
بالعودة إلى رياضايات تكون الأرض هي اقتران تكلفة , والأرتفاع في قاع الوادي يكون اقل نقطة في اقتران تكلفة

لنقم بالقاء نظرة على اقتران تكلفة الخاص بالانحدار :



Machine Learning for Humans

نستطيع رؤية ذالك على شكل اقتران بمتغيرين : ب0 و ب1 . المتغيرات س و ص تم تحديدها اثناء تدريب . نريد تقليل قيمة هذا الأقتران .



الاقتران هو ق(ب0 و ب1) = ز .لبداء عملية هبوط تدرج , يجب عليك اختيار قيم عشوائية لـب0 و ب1 .

وبعدها نقوم بأيجاد [المشتقة الجزئية](https://en.wikipedia.org/wiki/Partial_derivative) لأقتران الخسارة بنسبة لكل معطاة لـ(دز\دب0 , دز\دب1) , المشتقة الجزئية تدلنا على مدى تاثير المعطى عند زيادته او طرحه بقيمة صغيرة .

بعبارة أخرى ، كم ستزيد تقديراتك للدخل السنوي على افتراض أن التعليم العالي صفر (β0) يزيد من قيمة اقتران الخسارة (أي عدم دقة) النموذج الخاص بك؟ انت تريد

Machine Learning for Humans

اذهب في الاتجاه المعاكس بحيث ينتهي بك الأمر إلى المشي إلى الأسفل وتقليل قيمة اقتران الخسارة.

وبالمثل ، إذا قمت بزيادة التقدير الخاص بك كم كل سنة إضافية من التعليم يؤثر على الدخل(ب1) ، ما هو مقدار خسارة الخسارة هذه (ض)؟ إذا كان المشتق الجزئي دز /دب1 هو رقم سلبي ، ثم زيادة ب1 أمر جيد لأنه سيقلل من الخسارة الكلية. إذا إنه رقم موجب ، فأنت تريد تقليل ب1. إذا كان صفرًا ، فلا تغير ب1 لأنها يعني أنك وصلت إلى المستوى الأمثل

استمر في فعل ذلك حتى تصل إلى القاع ، بمعنى أنه تم التقليل من الخوارزمية وتقليل الخسارة. هناك الكثير من الحيل والحالات الاستثنائية خارج نطاق هذه السلسلة ، ولكن بشكل عام ، هذه هي الطريقة التي تجد بها المعلمات المثلى لنموذج حدودي.

اكداء

الأكداء : " شارلوك شرحك لما جرى كان مفصل بشدة "

تسوية : "لا تعقد الأمور شارلوك , سألكمك لكمة لكل كلمة زائدة"

المتغيرات-الفائقة (λ): "وهذه القوة التي سألكمك بها"

هناك مشكلة شائعة في تعلم الآلة وهي الأكداء : تعلم نموذج يشرح بيانات تدريب بشكل مطابق . ولكن لا يعمم على بيانات الأختبار . يحدث الأكداء عندما يتم يتعلم نموذج بشكل زائد من بيانات تدريب لدرجة إنه يجد خصويات لا تمثل الأنماط في العالم الواقعي . هذا يمثل مشكلة خصوصاً عندما يزيد تعقيد نموذجك . الأرخاء يحدث عندما يكون نموذجك بسيط بشكل مبالغ فيه لدرجة إنه لا يستطيع وصف الأنماط في بيانات تدريب.

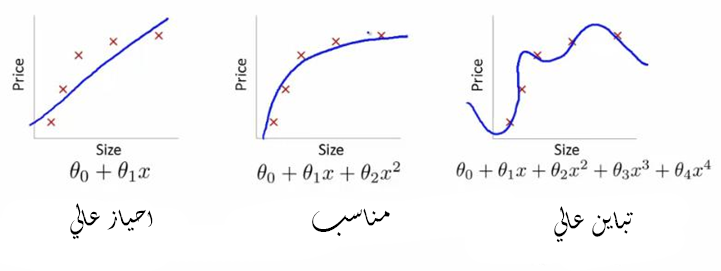
Machine Learning for Humans

مقايضة الأنحياز – التباين

**الأنحياز** : كمية الأخطىء ناتجة عن تقدير ظاهرة طبيعية بستخدام نموذج مبسط .

**التباين** :هو مقدار اختبار نموذجك لتغير الأخطاء بناء على تنوع في بينات تدريب وهو يعكس حساسية النموذج لخصائص مجموعة البيانات التي تم تدريبها عليها.

عندما يزداد نموذج التعقيد ويصبح أكثر مرونة (مرونة) ، ينخفض تحيزه (يقوم بعمل جيد في شرح بيانات التدريب) ، لكن الفرق يزداد (لا يعمم كذلك). في نهاية المطاف ، لكي يكون لديك نموذج جيد ، فأنت تحتاج إلى واحد مع انحياز منخفض و تباين منخفض.



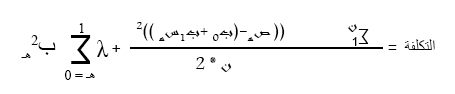
تذكر أن الشيء الوحيد الذي يهمنا هو كيفية اداء النموذج لبيانات الاختبار. أنت ترغب في التنبؤ رسائل البريد الإلكتروني التي سيتم تمييزها كرسائل غير مرغوب فيها قبل وضع علامة عليها ، وليس مجرد إنشاء نموذج دقيق بنسبة 100٪ عند إعادة تصنيف الرسائل الإلكترونية التي استخدمتها لإنشاء نفسها في المقام الأول. بعد فوات الأوان 20/20 - السؤال الحقيقي هو ما إذا كانت الدروس المستفادة ستساعد في المستقبل.

Machine Learning for Humans

نموذج الذي على اليمين ليس فيه خسارة لبيانات تدريب لانه يطابق كل نقاط البيانات . ولكن درس لا يعمم . سيخفق نموذج في توقع نقاط بيانات جديدة

طريقتان لمكافحة الاكداء:

1. ***استخدام المزيد من بيانات التدريب****. كلما كان لديك أكثر ، كلما كان من الصعب اكداء على البيانات عن طريق التعلم أكثر من أي مثال تدريب واحد.*
2. ***استخدم تسوية :*** *اضف ضريبة في افتران تكفلة , لبنائه نموذج يعطي الكثير من الأهمية لأي ميزة واحدة او لانه اخذ بالأعتبار الكثير من الميزات.*



شطر الأول من المجموع أعلاه هو اقتران الخسارة العادي. شطر الثاني هو عبارة عن مصطلح تسوية يضيف عقوبة لمعاملات بيتا كبيرة تعطي قوة تفسيرية أكبر لأي ميزة محددة. مع وجود هذين العنصرين ، يتوازن اقتران التكلفة الآن بين أولويتين: شرح بيانات التدريب ومنع هذا التفسير من أن يكون محددًا بشكل مفرط.

المعامل **لامدا** في عملية تسوية , عبارة عن معامل فائق : هو عبارة عن متغير عام الذي يمكن زيادته او انقاصه لضبط وتحصين اداء نموذج. كلما كبر لامدا زادت شدة عقوبة , لتحديد أفضل قيمة لـ (لامدا) ، يمكنك استخدام طريقة تسمى **التحقق المتبادل** والتي تتضمن الاحتفاظ بجزء من بيانات التدريب أثناء التدريب ، ومن ثم معرفة مدى جودة نموذجك في شرح جزء المحتفظ به. سنذهب إلى هذا بتعمق أكبر

لقد فعلتها !

Machine Learning for Humans

هذا ما تم تغطيته في هذا الفصل :

* كيف يسمح تعلم الغير مشرف بجعل الحواسيب قادرة على تعلم من البيانات بدون برمجة مسبقة
* وظائف تعليم المشرف الالانحدار و تصنيف
* **الالانحدار الخطي , اساس تعليم بمتغيرات**
* **تعلم المتغيرات و نحدار الميل**
* **الأكداء وتسوية**

في القسم التالي - الجزء 2.2: التعلم الخاضع للإشراف 2 - سنتحدث عن عنصرين مبدائين اساسياً التصنيف: الانحدار اللوجستي وآلات الدعم المتجه.

تمارين و موارد اضافية

2.1 أ – الحسر الخطي

للحصول على معالجة أكثر شمولاً الحسر الخطي ، اقرأ الفصول 1-3 من مقدمة في التعلم الإحصائي. الكتاب متاح مجانًا عبر الإنترنت وهو مورد ممتاز لفهم مفاهيم التعلم الآلي مع التدريبات المصاحبة.

للمزيد من تدريبات :

• العب مع مجموعة بيانات Boston Housing. يمكنك إما استخدام برنامج به واجهات رسومية لطيفة مثل Minitab و Excel أو القيام بذلك بالطريقة الصعبة (ولكن الأكثر إفادة) باستخدام Python أو R.

* جرب يدك في تحدي Kaggle ، على سبيل المثال توقع أسعار المساكن ، وانظر كيف تعامل الآخرون مع المشكلة بعد أن حاولت ذلك بنفسك.

2.1 ب - تطبيق هبوط المتدرج

. لتنفيذ الانحدار في بايثون فعليًا ، راجع هذا البرنامج التعليمي. وهنا وصف أكثر دقة من الناحية الحسابية لنفس المفاهيم.

من الناحية العملية ، نادرًا ما تحتاج إلى تنفيذ هبوط التدرج من الصفر ، لكن فهم كيفية عمله خلف الكواليس سيسمح لك باستخدامه بشكل أكثر فاعلية وفهم سبب تعطل الأشياء عند حدوثها.

Machine Learning for Humans

الجزء ثاني الباب ثاني تعلم المشرف



تصنيف بستخدام الحسر الوغرثمي و تأيد متجهة الآلة (أمل)

تصنيف : توقع تسمية

*هل هذا البريد الإلكتروني غير مرغوب فيه أم لا؟ هل سيقوم المقترض بسداد قرضه؟ هل سينقر المستخدمون على الإعلان أم لا؟ من هو هذا الشخص في صورتك على الفيسبوك؟*

تصنيف هو تبناء بتسمية الهدف ص و تصنيف هو مشكلة تعين ملاحظات , للفئة التي تكون مشابه لهذه الملاحظات . على أساس نموذج تصنيف مبني من بيانات التدريب المسمى.

تعتمد دقة التصنيفات الخاصة بك على فعالية الخوارزمية التي تختارها ، وكيفية تطبيقها ، ومقدار بيانات التدريب المفيدة التي لديك.



الحسر الوغرثمي : صفر او واحد ؟

الحسر الوغرثمي : هو طريقة **لتصنيف** . يقوم نموذج بأخراج نسبة كون المدخل (ص) ينتمي **لفئة معينة**

مثال جيد على تصنيف هو اقرار اذ كان المقدم على قرض محتال ام لا

في نهاية , المُقرض يريد إن يعرف هل يُقرض المقترض أم لا , ولدى المٌقرض نسبة من تسامح او الخطا في كون المقترض محتال . في هذه الحالة الحسر الوغرثمي يحسب احتمالية بين 0% و 100% كون المقترض محتال . وبناء على هذه الأحتمالات نستطيع وضع الحد الأدنى الذي يقرض من يتجاوزه ويرفض طلب كل من لم يتجاوزه او يتم تعليم طلب للمراجعه .

الحسر الوغرثمي يستخدم في **تصنيف الثنائي** . حيث يكون هناك فئتان . وتذكر ان تصنيف قد يكون على اكثر من فئة مثل :(تعين خانة مكتوبة باليد فئة بين 0 و 9 أو استخدام تعرف على الوجوه لتعرف على الأصدقاء في صورة (

**هل استطيع استخدام طريقة القيمة الصغرى لمجموع المربعات ؟**

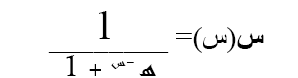
لا . اذ قمت بتدريب نموذج الانحدار خطي على مجموعة امثلة حيث ص = 0 أو ص = 1 . ستنتهي بإن هناك احتمالية بكون القيمة المتوقعة اقل من 0 او اكبر من 1 وهذا غير منطقي . عوضاً عن ذالك نستخدم نموذج لوغرثمي والذي يقوم بتعين قيمة بين 0% او 100% بكون القيمة الجديدة تنتمي لفئة معينة.

**الجانب الرياضي؟**

ملاحظة : الرياضيات في هذا القسم مثيرة للأهتمام ولكن قد تكون على جانب تقني .

قم بالقاء نظرة سريعة إذ كنت مهتم بالمفاهيم الفكرية . نموذج الحسر الوغرثمي هو تعديل على نموذج الحسر الخطي لا تتردد في تصفحها إذا كنت مهتمًا بالمفاهيم عالية المستوى. إن نموذج logit هو تعديل الانحدار الخطي الذي يضمن إخراج إحتمال بين 0 و 1 من خلال تطبيق وظيفة sigmoid ، والتي ، عند رسمها ، تشبه المنحنى المميز على شكل S الذي سترونه في وقت لاحق.

Machine Learning for Humans



اقتران sigmoid: يقم باخراج قيمة بين 0 و 1

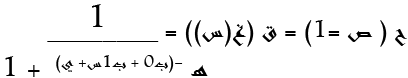
تذكر الأقتران الخطي من مثالنا في الحسر الخطي, والذي سنسميه غ(س) وبما اننا سنقوم بستخدام اقتران مركب:



لحل مشكلة كون ناتج اصغر من صفر او اكبر من واحد , سنقوم بتعريف اقتران جديد ق(غ(س)) الذي سيقوم عبر تغير ناتج الأقتران الخطي إلى قيمة بين [0,1] , هل تفكر بأقتران يستطيع فعل ذالك ؟

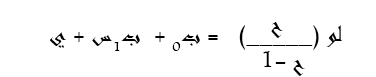
هل تفكر في وظيفة السيجمويد؟ بام! المعزوفة! أنت على حق.

لذلك نقوم بتوصيل g (x) بوظيفة السيجمويد الموضحة أعلاه ، مما يؤدي إلى وظيفة من وظيفتنا الأصلية (نعم ، الأشياء تحصل على ميتا) والتي تنتج احتمال بين 0 و 1:



بمعنى آخر ، نحن نحسب احتمال أن يكون مثال التدريب ينتمي إلى فئة معينة: P (Y = 1).

هنا قمنا بمعزل p ، احتمال أن Y = 1 ، على الجانب الأيسر للمعادلة. إذا كنا نرغب في حل للحصول على نظيف نظيف β0 + β1x + ϵ على الجانب الأيمن حتى نتمكن من تفسير معاملات بيتا التي سنتعلمها بشكل مباشر ، فسننتهي بدلاً من ذلك باستخدام نسبة تسجيل الأرجحية أو تسجيل الدخول ، على الجانب الأيسر - ومن هنا جاءت تسميته "نموذج logit":



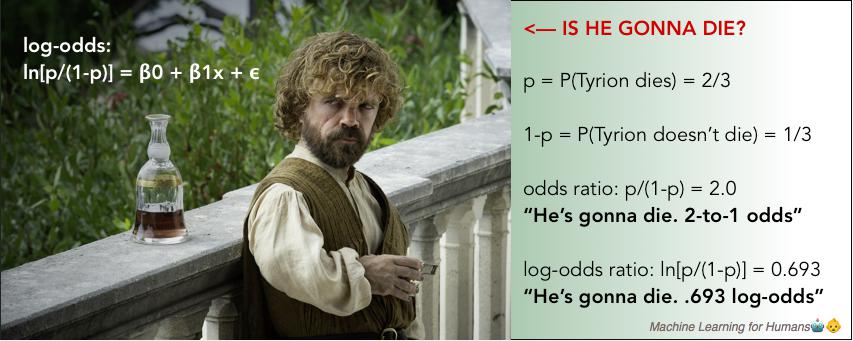
Machine Learning for Humans

نسبة اللوغاريتمات - الأرجحية هي ببساطة السجل الطبيعي لنسبة الأرجحية ، p / (1-p) ، والتي تظهر في المحادثات اليومية:

*"هآي , ما هو احتمال إن تيرون لانستر سيموت في هذا الموسم "”*

*(كارثة الموسم ثامن جعلت هذا المرجع موجعا ☹)*

*"هم. من المؤكد أنه من المرجح أن يحدث مرتين أكثر من ذلك. 2 إلى 1 خلاف. بالتأكيد ، قد يبدو من الأهمية بمكان أن يُقتل ، لكننا رأينا جميعاً ما فعلوه مع نيد ستارك ... ".*



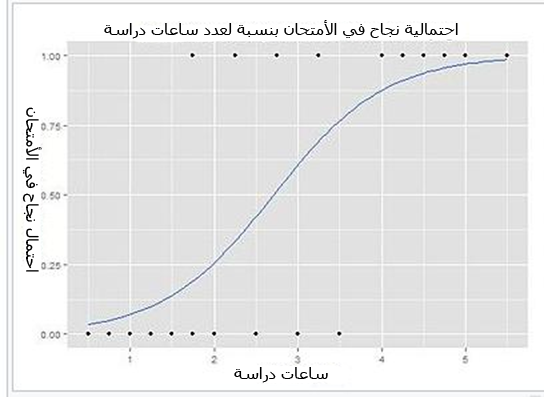
لاحظ أنه في نموذج اللوغاريثمي، يمثل ب1 الآن معدل التغير في نسبة الأرجحية - الوغرثمية مع تغير س. بعبارة أخرى ، إنه "منحدر من الاحتمالات" وليس "منحدر الاحتمال".

قد يكون الأرجحية - الوغرثمية غير واضح إلى حد ما ، ولكن الأمر يستحق الفهم لأنه سيظهر مرة أخرى عند تفسير مخرجات الشبكات العصبية التي تؤدي مهام التصنيف.

Machine Learning for Humans

**بستخدام ناتج نموذج الحسر الوغرثمي نستطيع اتخاذ القرارات**

مخرج نموذج الانحدار الوغرثمي من الأعلى يبدو على شكل حرف [S]ح(ص=1) بنائاً على قيمة س.



Source: [Wikipedia](https://en.wikipedia.org/wiki/Logistic_regression)

. للتنبؤ بالعلامة "ص" البريد المزعج / ليس البريد المزعج ، السرطان / ليس السرطان ، الاحتيال / عدم الاحتيال ، وما إلى ذلك - يتعين عليك تعيين **عتبة** ، للحصول على نتيجة إيجابية. على سبيل المثال: "إذا اعتقد نموذجنا أن احتمال أن تكون هذه الرسالة الإلكترونية غير مرغوب فيها أعلى من 70٪ ، فقم بتمييزها كرسالة غير مرغوب فيها. وإلا ، لا تفعل ذلك"

يعتمد عتبة على التسامح الخاص بك للحصول على **إيجابيات كاذبة** مقابل **السلبيات الكاذبة**. إذا كنت تشخص الإصابة بالسرطان ، فلديك قدر كبير من التسامح مع السلبيات الكاذبة ، لأنه حتى إذا كانت هناك فرصة ضئيلة جدًا للمريض بالسرطان ، فستحتاج إلى إجراء المزيد من الاختبارات للتأكد من ذلك. لذا ، يمكنك تعيين عتبة منخفضة جدًا للحصول على نتيجة إيجابية.

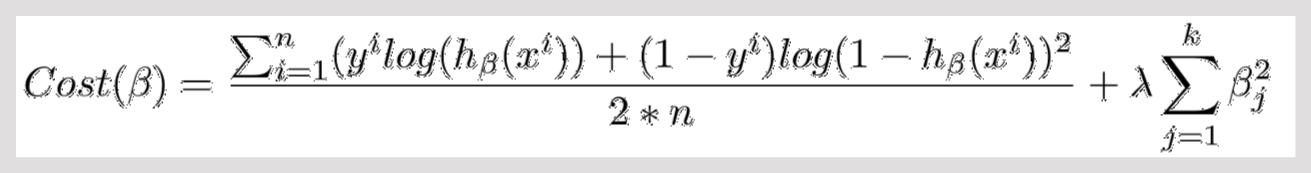
Machine Learning for Humans

في حالة طلبات القروض الاحتيالية ، من ناحية أخرى ، قد يكون التسامح بشأن الإيجابيات الزائفة أعلى ، خاصة بالنسبة للقروض الأصغر ، نظرًا لأن المزيد من التدقيق أمر مكلف ، وقد لا يستحق قرض صغير تكاليف التشغيل الإضافية والاحتكاك في حالة عدم الاحتيال المتقدمين الذين يتم الإبلاغ عن مزيد من المعالجة.

**تقليل الخسارة مع االانحدار المنطقي**

كما هو الحال في حالة الانحدار الخطي ، نستخدم أصل التدرج لمعرفة معلمات بيتا التي تقلل من الخسارة.

في الانحدار الوغرثمي ، يعتبر اقتران التكلفة قياسًا لعدد المرات التي تنبأت فيها 1 عندما كانت الإجابة الصحيحة 0 ، أو العكس. يوجد أدناه اقتران تكلفة منتظمة **مستوي** تمامًا مثل تلك التي قمنا بها للتحول الخطي.



لا داعي للذعر عندما ترى معادلة طويلة كهذه! قسّمها إلى أجزاء وفكر في ما يجري في كل جزء من الناحية المفاهيمية. ثم ستبدأ التفاصيل في أن تكون منطقية.

الجزء الأول هو فقدان البيانات ، أي مقدار التباين الموجود بين تنبؤات النموذج والواقع. الجزء الثاني هو فقدان التسوية ، بمعنى ما مقدار العقوبة على النموذج لوجود متغيرات تساهم بشكل كبير في سمات معينة ،( تذكر هذا يمنع الأكداء)

سنقلل من اقتران تكلفة بستخدام هبوط تدرج , كما هو موضح في صورة فوق , ولقد بنينا نموذج الانحدار لوغرثمي

Machine Learning for Humans

**تأيد متجهة الآلة (امل)**

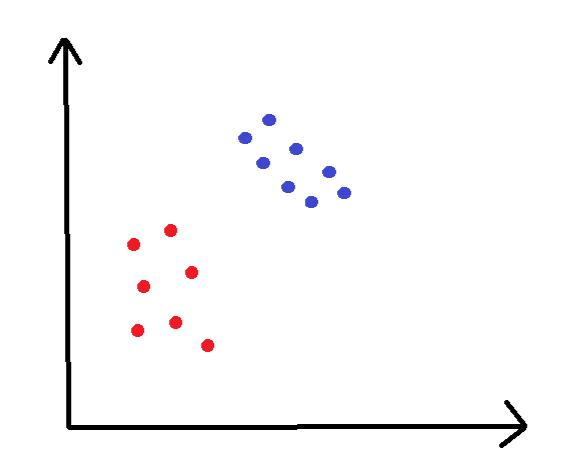
*"نحن في غرفة مليئة بكرات زجاجية مرة أخرى. لماذا نحن دائما في غرفة مليئة بكرات زجاجية ؟ كان من الممكن أن أقسم أننا فقدناهم بالفعل ".*

**امل** ستكون اخر نموذج بمعطيات سنقوم بتغطيته . في العادة تقوم بحل نفس المشكلات التي يحلها الالانحدار الوغرثمي – تصنيف بين فئتان – وبينتج عنها نتائج مشابه , ويجدر فهمها لانها مبنية على الهندسة عوضاً عن الأحتمالات.

بعض الأمثلة على المشكلات التي يمكن أن تحلها بستخدام تأيد متجه الآلة :

* هل هذه صورة لقطة أو كلب؟
* هل هذه المراجعة إيجابية أم سلبية؟
* هل نقاط في هذا المستوى زرقاء ام حمراء؟

سنقوم بستخدام المثال ثالث لتوضيح كيف تعمل (أمل) , المشكلة هذه تسمى مشكلة للعبة لانها ليست حقيقة .



في هذا المثال ، لدينا نقاط في فضاء ثنائية الأبعاد إما حمراء أو زرقاء ، ونود أن نفصل بين الاثنين.

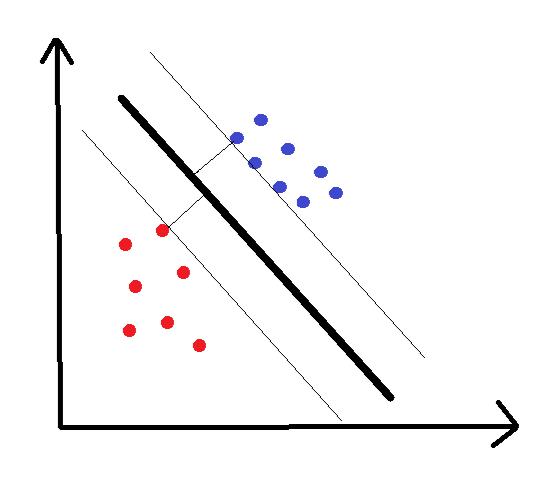
بيانات تدريب تم رسمها . نود إن نصنف نقطة جديدة غير مصنفة . لفعل ذالك تقوم (أمل) برسم خط فاصل (وفي حال كان لدينا فضاء ثلاثي الابعاد ترسم مسطح) لفصل الفضاء إلى منطقة حمراء و منطقة زرقاء . تستطيع تخيل الخط الفاصل الذي سيفصل المجموعتان .

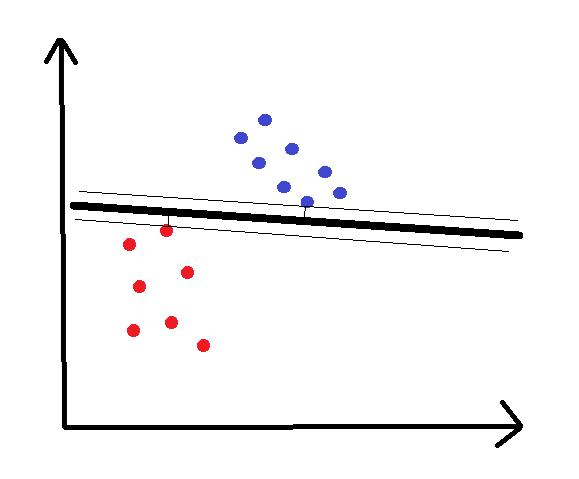
Machine Learning for Humans

منطقة زرقاء . تستطيع تخيل الخط الفاصل الذي سيفصل المجموعتان

كيف , بتحديد نختار اين نرسم الخط ؟

لدينا مثالان لهذا الخط:





نأمل أنك لديك الحدس في أن الخط الأول متفوق. تسمى المسافة إلى أقرب نقطة على جانبي الخط الهامش ، وتحاول (أمل) زيادة الهامش. يمكنك التفكير في الأمر كمساحة أمان: كلما كبرت المساحة ، قل احتمال تعرض النقاط المزعجة لسوء التصنيف.

Machine Learning for Humans

بناء على هذا التفسير القصير ، تأتي بعض الأسئلة الكبيرة.

**1. رياضيات خلف هذا**

نريد العثور على المسطح الأمثل (خط , في مثالنا ثنائي الأبعاد) يحتاج هذا المسطح إلى :

(1) فصل البيانات بشكل نظيف ، مع وجود نقاط زرقاء على جانب واحد من الخط ونقاط حمراء على الجانب الآخر

(2) زيادة الهامش. هذه مشكلة تحسين. الحل يجب أن يحترم القيد (1) مع زيادة الهامش كما هو مطلوب في (2)

ستكون نسخة البشرية لحل هذه المشكلة هو اخذ مسطرة والاستمرار في تجريب خطوط مختلفة تفصل بين جميع النقاط حتى تحصل على الحد الذي يزيد من الهامش.

اتضح أن [هناك طريقة رياضية](https://en.wikipedia.org/wiki/Lagrange_multiplier) لتحقيق هذا الحد الأقصى ، ولكن التفاصيل خارج نطاقنا. لاستكشافها بشكل أكبر ، إليك [محاضرة](https://www.youtube.com/watch?v=_PwhiWxHK8o) فيديو توضح كيفية عملها باستخدام ميزة.

. يتم تحديد المستوى الفائق للحل الذي ينتهي به الأمر فيما يتعلق بموضعه فيما يتعلق ببعض قيم x\_i ، والتي تسمى متجهات الدعم ، وعادة ما تكون الأقرب إلى المستوى الفائق.

**2.ماذا يحدث إذا لم تتمكن من فصل البيانات بشكل كامل؟**

يوجد منهجان لتعمل مع هذه المشكلة

*2.1 نقوم بتخفيف شدة " الفصل"*

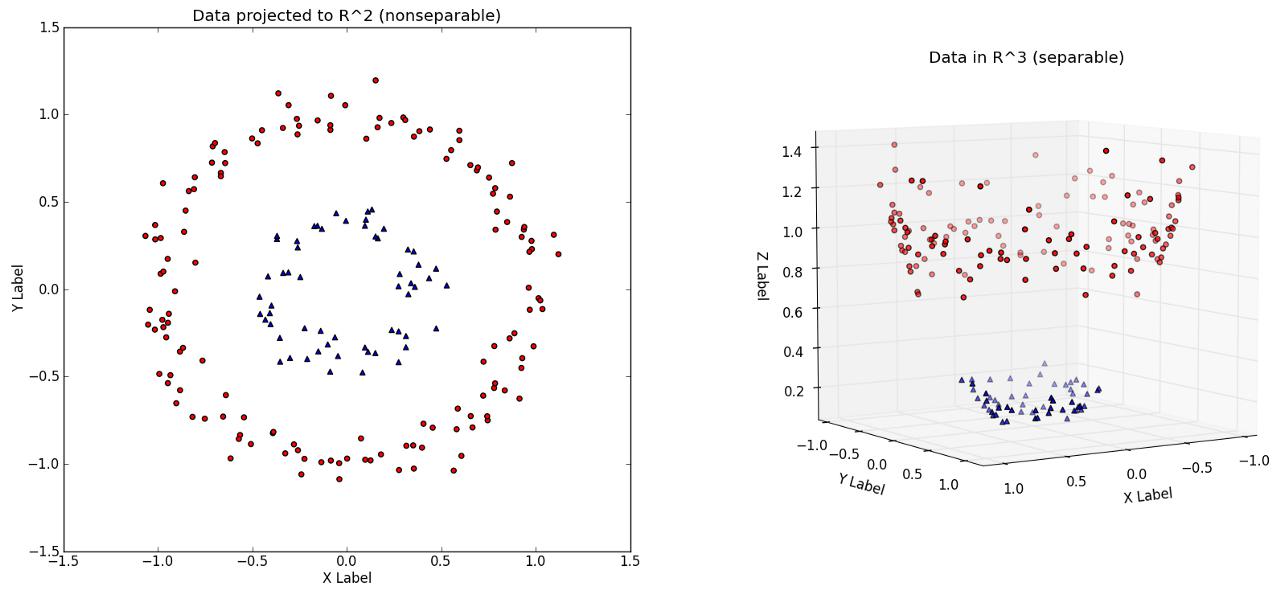
نحن نسمح ببعض الأخطاء ، بمعنى أننا نسمح ببعض النقاط الزرقاء في المنطقة الحمراء أو بعض النقاط الحمراء في المنطقة الزرقاء. نقوم بذلك عن طريق إضافة تكلفة C لأمثلة مصنفة بشكل خاطئ في اقتران تكلفة. في الأساس ، نقول أنه من المقبول ولكن من المكلف تصنيف نقطة ما بشكل خاطئ.

*2.2 نقوم برفع نقاط إلى ابعاد عليا*

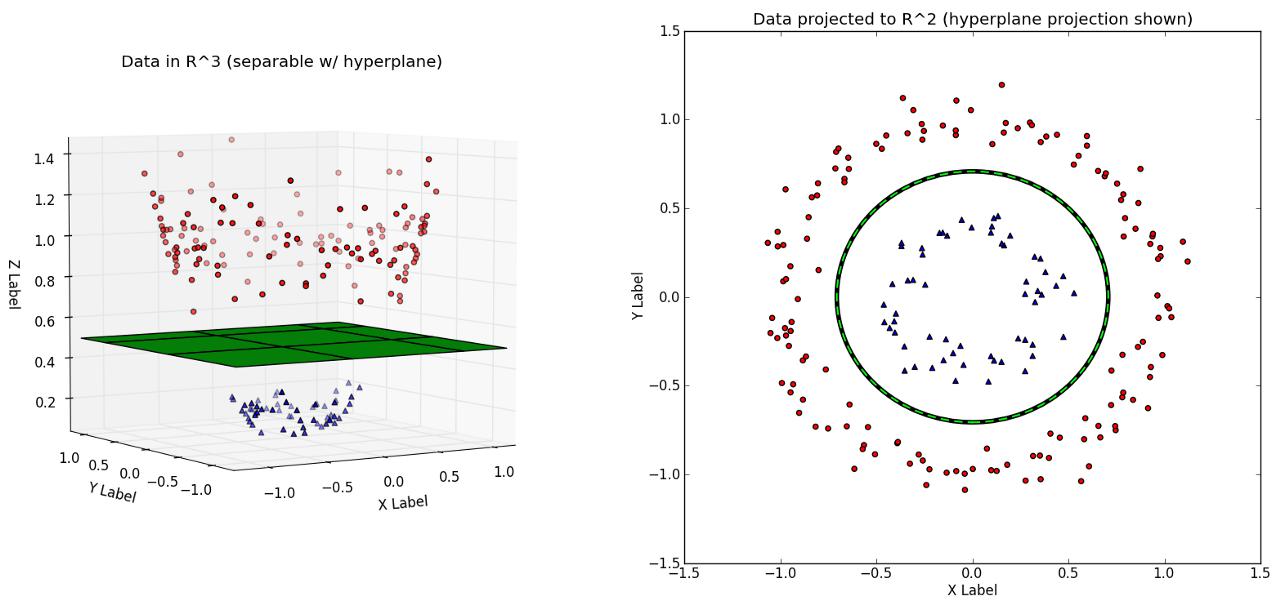
مصنفات غير خطية عن طريق زيادة عدد الأبعاد ، على سبيل المثال تضمين x² ، x³ ، حتى cos (x) ، إلخ. فجأة ، يكون لديك حدود يمكن أن تبدو متعرجة أكثر عندما نعيدها إلى التمثيل الأقل بعدًا.

Machine Learning for Humans

بشكل بدهي ، هذا يشبه وجود كرات أحمر وأزرق ملقى على الأرض بحيث لا يمكن فصله بشكل نظيف عن طريق خط - ولكن إذا تمكنت من جعل كل الرخام الأحمر يطف عن الأرض بالطريقة الصحيحة تمامًا ، يمكنك رسم مسطح يفصلهما. ثم سمحت لهم بالعودة إلى الأرض لمعرفة أين تتوقف الكرات زرقاء وتبدأ الحمراء.



مجموعة بيانات غير قابلة للفصل في فضاء ثنائي الأبعاد **ر2** ، ونفس مجموعة البيانات تم تعيينها على أبعاد ثنية ذات البعد الثالث هو (المصدر: [http://www.eric-kim.net/eric-kim-net/posts/1 /kernel\_trick.html](http://www.eric-kim.net/eric-kim-net/posts/1%20/kernel_trick.html))



تظهر حدود القرار باللون الأخضر ، أولاً في الفضاء ثلاثي الأبعاد (على اليسار) ، ثم مرة أخرى في الفضاء ثنائي الأبعاد (على اليمين). نفس مصدر الصورة السابقة.

، يتم استخدام (أمل) للتصنيف مع فئتين. يحاولون العثور على طائرة تفصل بين الفئتين بشكل نظيف. عندما لا يكون هذا ممكنا ، إما أن نخفف تعريف "منفصل" ، أو نطرح البيانات في أبعاد أعلى حتى نتمكن من فصل البيانات بشكل واضح.

نجحنا!

في هذا الفصل لقد مررنا بـ:

* مهمة التنصيف من تعليم المشرف
* طريقتان للتصنيف: الانحدار الوغرثمي تأيد متجه الآلة (أمل)
* المفاهيم المتكررة: اقتران السيجمويد ، الاحتمالات اللوغاريثمية ، والإيجابيات الكاذبة مقابل السلبيات الزائفة ،

في الجزء 2.3: التعلم المشرف 3 ، سننتقل إلى التعلم المشرف غير المعياري ، حيث تكون الأفكار الكامنة وراء الخوارزميات بديهية جدًا والأداء ممتاز لأنواع معينة من المشكلات ، ولكن قد يكون من الصعب تفسير النماذج.

تدريبات و مراجع اضافية

**2.2 أ – الحسر الوغرثمي**

مدرسة البيانات لديها دليل ممتاز متعمق للانحدار اللوجستي. سنستمر أيضًا في إحالتك إلى مقدمة حول التعلم الإحصائي. انظر الفصل 4 حول الانحدار اللوجستي ، والفصل 9 حول آلات المتجهات الداعمة.

لتنفيذ الانحدار **الوغرثمي** ، نوصي بالعمل على مجموعة المشكلات هذه. يجب عليك التسجيل في الموقع للعمل من خلاله للأسف. هذه هي الحياة.

**2.2 ب - تأيد متجهة الآلة**

للتعمق في الرياضيات الكامنة خلف (أ/مل) ، شاهد محاضرة البروفيسور باتريك وينستون من معهد ماساتشوستس للتكنولوجيا 6.034: الذكاء الاصطناعي. وتحقق من هذا البرنامج التعليمي للعمل من خلال تطبيق.

Machine Learning for Humans

الفصل ثاني الباب ثالث : تعليم المشرف



*نماذج بدون معاملات : الجار الأقرب لـ(ك) , اشجار القرار , الغابات العشوائة . مقدمة الى تصديق متقاطع , ضبط المتغيرات الفائقة و تجميع نماذج*

المتعلمات بدون معملات مدخلات

على نقيض طرق التي غطيناها حتى الأن —  الحسر الخطي , الحسر الوغرثمي و تأيد متجهة الآلة حيث يكون نموذج معرف مسبقاً —  المتعلمات بدون مدخلات لا تمتلك نموذج معرف مسبقاً . نقوم بتخمين حول شكل الأقتران ق الذي نحاول تعلم منه قبل تدريب نموذج, كما فعلنا سابقاً بستخدام الالانحدار الخطي ولكن بنية نموذج تكون مقررة من قبل البيانات فقط.

هذه تنماذج تكون اكثر مرونة وتقبلاً لشكل البيانات ولكن هذا يأتي على حساب قابلية نموذج لتفسير

الجار الأقرب لـ(ك) : (جارك)

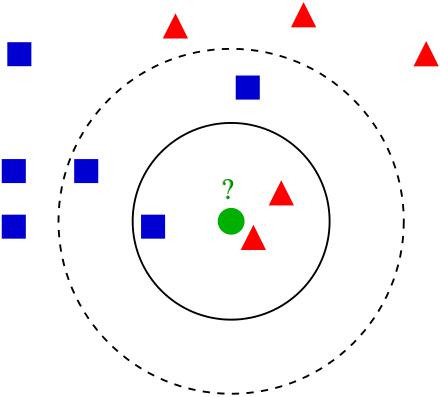
“انت متوسط (ك) من افضل اصدقائك”

في الوهلة الأولى تبدو خورزمية (جارك) بسيطة جداً لأن تكون من خورزميات تعلم الآلة . المبداء هو تسمية نقطة الأختبار س عبر ايجاد الوسيط او المنوال لـأقرب نقاط بيانات مسمية .

قم بالقاء نظرة على صورة في الأسفل . لنفترض إنك تريد معرفة إذ كانت دائرة الخضراء . مثلث احمر او مربع ازرق . كيف تفعلها ؟

Machine Learning for Humans

تستطيع المحاولة عبر إن تأتي بمعادلة تقوم بأخذ موقع دائرة الخضراء وتقوم بأعطاء توقع مناسب . او تستطيع اخذ اقرب ثلاث جيران وتعرف إن دائرة الخضراء على الأرجح مثلث احمر . وتستطيع توسيع المحيط لتشمل خمس جيران , وفي ذالك تكون دائرة الخضراء مربع ازرق على الأرجح. ( 3\5 من اقرب الجيران هم مربعات زرقاء لذالك تكون دائرة الخضراء على الأرجح مربع ازرق عندما تكون (ك) = 5 )



توضيح جارك (ك = 1,3,5) . لتصنيف الدائرة الخضراء الغامضة (س) أعلاه ، انظر إلى أقرب جار منفرد لها ، وهو "المثلث الأحمر". لذلك ، نحن نعتقد أن صَ = "المثلث الأحمر". مع (ك = 3) ، انظر إلى الجيران الثلاثة المجاورة: وضع هذه العناصر هو "المثلث الأحمر". مرة أخرى ، لذا , مع (ك = 5) ، نأخذ وضع أقرب 5 جيران بدلاً من ذلك. الآن ، لاحظ أن صَ اصبحت "المربع الأزرق". صورة من ويكيبيديا.

هذه هي (جارك) تقوم بأخذ (ك) من اقرب نقاط البيانات وتأخذ متوسط القيم إذ كان المتغير متصل (مثل سعر المنزل) او المنوال إذ كان المتغير عبارة منفصل (مثل فئة (قط او كلب)).

إذ كنت تريد معرفة سعر منزل مجهول , تستطيع إن تأخذ متوسط بعض المنازل القريبة جغرفياً , وفي العادة تكون قد حصلت على تخمين جيد وقد يتفوق هذا تخمين على نموذج الانحدار بمدخلات الذي صنع من قبل اقتصادي قدر إن معاملات نموذج # عدد الغرف \ عدد الحمامات , القرب من مدرسة , المسافة للمواصلات العامة , الخ.

كيف تستخدم (جارك) لتوقع سعر منزل:

1. خزن البيانات في مصفوفة (س) تحتوي على صفات مثل الموقع جغرافي , الحي , عدد غرف نوم , المساحة , المسافة للمواصلات العامة ..الخ و المصفوفة (ص) التي تحتوي على الأسعار
2. رتب المنازل في مجموعة البيانات حسب مقدار تشابه مع المنزل الذي يراد معرفة سعره بنائاً على الصفات في (س) (سيتم تعريف تشابه لاحقاً)
3. خذ الوسيط لأقرب (ك) من المنازل , الوسيط سيكون سعر المنزل

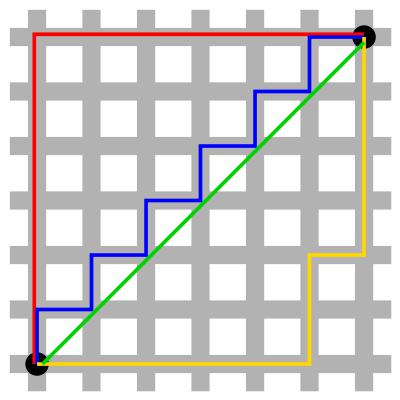
واقعة إن (جارك) لا تحتاج إلى معاملات معرفة مسبقاً تجعلها مناسبة للمواقف حيث تكون العلاقة بين المتغرات معقدة جداً كي يتم تعبير عنها بشكل خطي

مقياس القرب : تعريف وحساب الـ(قرب)

العثور على "أقرب الجيران"؟ كيف تحدد رياضيا أي من المربعات الزرقاء والمثلثات الحمراء في المثال أعلاه هي الأقرب إلى الدائرة الخضراء ؟

Machine Learning for Humans

اكثر طريقة مباشرة هي عبر (المسافة اليقلودنية) وطريقة اخر هي مسافة مانهاتن وهي طريقة التي تحتاج لأقل قدر من الحسابات



الخط الأخضر = المسافة الإقليدية. الخط الأزرق = مسافة مانهاتن. المصدر: ويكيبيديا

هل تتذكر نظرية فيثاغورس لإيجاد طول وتر المثلث القائم؟



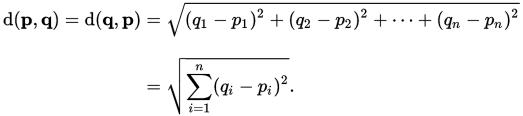
ج = طول الوتر (الخط الأخضر أعلاه). أ و ب = طول الضلعين الآخرين بزاوية قائمة (خطوط حمراء أعلاه).

Machine Learning for Humans

بالحل بدلالة c ، نجد طول الوتر بأخذ الجذر التربيعي لمجموع أطوال التربيعية لـ a و b ، حيث a و b ضلعان متعامدان للمثلث (أي أنهما بزاوية 90 درجة من أحدهما الآخر ، يسير في اتجاهات متعامدة في الفضاء).



فكرة إيجاد طول الوتر المعطى في اتجاهين متعامدين تعمم على العديد من الأبعاد ، وهذه هي الطريقة التي نشتق بها معادلة المسافة الإقليدية d (p ، q) بين النقطتين p و q في الفضاء ذي الأبعاد ن:



. صيغة للمسافة الإقليدية ، مشتقة من نظرية فيثاغورس.

باستخدام هذه الصيغة ، يمكنك حساب قرب جميع نقاط بيانات التدريب من نقطة البيانات التي تحاول تصنيفها ، واتخاذ متوسط / وضع k الأقرب جيرانًا لعمل توقعك.

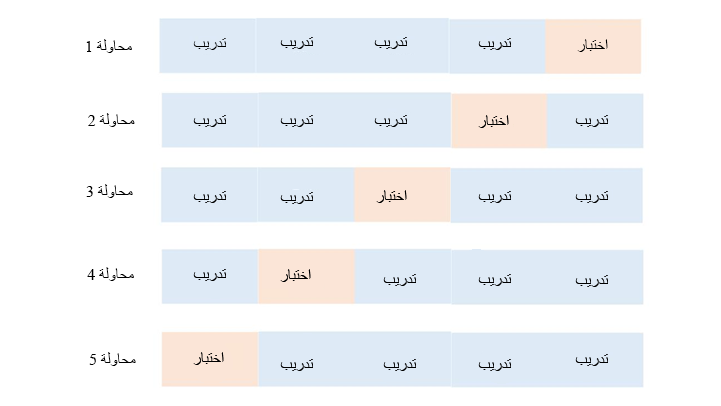
لن تحتاج عادةً إلى حساب أي مقاييس للمسافة يدويًا - يكشف بحث سريع عن الاقترانات الجاهزة في NumPy أو SciPy والتي ستقوم بذلك نيابةً عنك ،على سبيل المثال euclidean\_dist = numpy.linalg.norm (p-q) - لكن من الممتع أن نرى كيف أن مفاهيم الهندسة من الصف الثامن في نهاية المطاف مفيدة لبناء نماذج تعلم الآلة اليوم!

Machine Learning for Humans

**اختيار (ك): ضبط المعلمات الفائقة مع التحقق المتقاطع**

لتحديد قيمة (ك) المراد استخدامها ، يمكنك اختبار نماذج (جارك) مختلفة باستخدام قيم مختلفة لـ k **تصديق متقاطع**:

1. قم بتقسيم بيانات التدريب إلى شرائح ، وقم بتدريب النموذج الخاص بك على جميع القطاعات باستثناء جزء واحد ؛ استخدام الجزء المعلق كبيانات "اختبار".
2. تعرف على كيفية أداء نموذجك من خلال مقارنة تنبؤات النموذج (ŷ) بالقيم الفعلية لبيانات الاختبار (ص).
3. اختر أيهما ينتج عنه أقل خطأ ، في المتوسط ، عبر جميع التكرارات.



Cross-validation illustrated. The number of splits and iterations can be varied.

Machine Learning for Humans

أعلى (**ك**) يمنع الاكداء

تساعد القيم الأعلى لـ (**ك**) في معالجة التجاوز ، ولكن إذا كانت قيمة k عالية جدًا ، فسيكون نموذجك متحيزًا للغاية وغير مرن. لنأخذ مثالًا متطرفًا: إذا كان (**ن**) = (**ك**) (العدد الإجمالي لنقاط البيانات) ، فإن النموذج سيصنف جميع بيانات الاختبار على أنها متوسط أو وضع بيانات التدريب.

إذا كان الحيوان الأكثر شيوعًا في مجموعة بيانات الحيوانات هو القط الاسكتلندي ، فإن (جارك)مع تعيين (**ك**) على (**ن**)(رقم ملاحظات التدريب) سيتنبأ حينئذٍ بأن كل حيوان آخر في العالم هو أيضًا قطط اسكتلندية. وهو ، في رأي فيشال ، سيكون رائعًا. سامر لا يوافق.

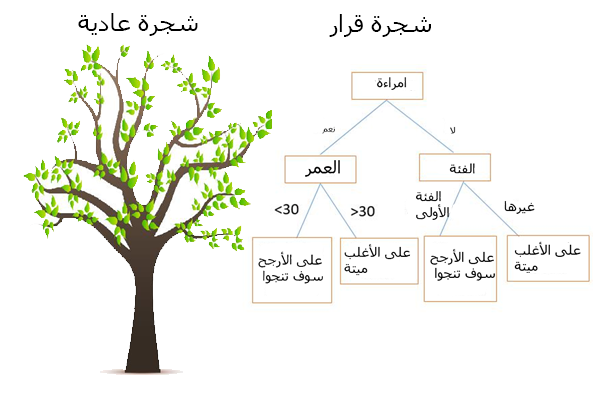


Machine Learning for Humans

اين تستخدم (جارك) في الواقع؟

* **التصنيف: الكشف عن الأحتيال.** يمكن تحديث النموذج على الفور تقريبًا بأمثلة تدريب جديدة نظرًا لأنك تقوم فقط بتخزين المزيد من نقاط البيانات ، مما يتيح التكيف السريع مع أساليب الاحتيال الجديدة.
* **الحسر**: التنبؤ بأسعار المساكن. في التنبؤ بأسعار الإسكان ، حرفيًا كونك "جارًا قريبًا" هو في الواقع مؤشر جيد على أن تكون متشابهًا في السعر. (جارك) مفيد في المجالات التي يكون فيها القرب المادي مهمًا.
* **إدخال بيانات التدريب المفقودة**. إذا كان أحد الأعمدة في ملف .csv الخاص بك يحتوي على الكثير من القيم المفقودة ، فيمكنك إدراج البيانات عن طريق أخذ الوسط أو الوضع. يمكن أن يوفر لك (جارك) تخمينًا أكثر دقة إلى حد ما عند كل قيمة مفقودة.

اشجار القرارات و الغابات العشوائية



Machine Learning for Humans

اول انقسام يحدث عند الجذر , حيث يشابه هذا الأنقسام سؤال الأول في للعبة العشرين سؤال . يتوجب عليك فصل البيانات بشكل كلي . لكي تزيد من كمية المعلومات المكتسبة

إذا سألك احد أصدقائك (انا أفكر في أسم , أسألني عشرين سؤال نعم أو لا لتحزر ما هو) , هل هو بطاطا , سيجب بـ(لا) ولن تكتسب اي معلومات جديدة , إلا أذ كان صديقك يفكر في البطاطا دائما

بدلاً من ذلك ، قد يكون السؤال مثل "هل هو كائن؟" أكثر منطقية.

هذا هو نوع من الأمثلة على كيفية فرز المرضى المستشفيات أو نهج التشخيص التفريقي. يسألون بعض الأسئلة ويتحققون من بعض العناصر الأساسية لتحديد ما إذا كنت ستموت أو شيء ما. إنهم لا يبدأون بإجراء الخزعة للتأكد من إصابتك بسرطان البنكرياس بمجرد أن تدخل من الباب.

هناك طرق لتحديد مقدار المعلومات ، بحيث يمكنك بشكل أساسي تقييم كل انقسام محتمل من بيانات التدريب وزيادة الحصول على المعلومات لكل قسم. وبهذه الطريقة يمكنك التنبؤ بكل تسمية أو قيمة بأكبر قدر ممكن من الكفاءة.

الآن ، دعونا نلقي نظرة على مجموعة بيانات معينة ونتحدث عن كيفية اختيار تقسيمات.

**مجموعة بيانات تيتانيك**

تحتوي [Kaggle] على مجموعة بيانات تيتانيك تستخدم في الكثير من عمليات التعلم الآلي. عندما غرقت تيتانيك ، مات 1502 من أصل 2224 من الركاب والطاقم. على الرغم من وجود بعض الحظ ، إلا أن النساء والأطفال والطبقة العليا كانوا أكثر عرضة للبقاء على قيد الحياة.

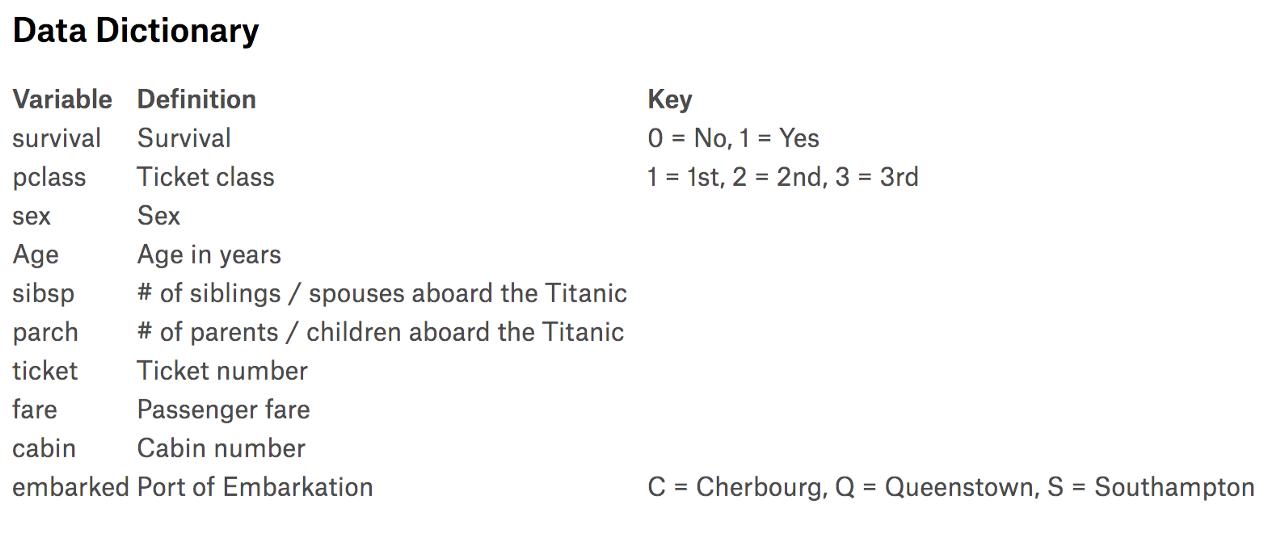
إذا نظرت إلى الوراء في شجرة القرار أعلاه ، سترى أنها تعكس نوعًا ما هذا التباين بين الجنسين والعمر والطبقة.

Machine Learning for Humans

**اختيار تقسيمات في شجرة القرار**

الفوضى هو مقدار الاضطراب في مجموعة (تقاس بمؤشر [جيني](https://en.wikipedia.org/wiki/Gini_coefficient) أو [الانتروبيا](https://en.wikipedia.org/wiki/Cross_entropy)). إذا كانت القيم مختلطة بالفعل ، فهناك الكثير من الفوضى ؛ إذا كنت تستطيع تقسيم القيم بشكل نظيف ، فليس هناك فوضى. لكل انقسام في نقطة الأصل ، تريد أن تكون نقطة الفرعية نقية قدر الإمكان - تقلل من الفوضى. على سبيل المثال ، في التايتانيك ، يعد طبقة الاجتماعي محددًا كبيرًا للبقاء على قيد الحياة ، لذلك فمن المنطقي أن يتم استخدام هذه الميزة في الفصل الأول لأنها الميزة التي تؤدي إلى الحصول على أكبر قدر من المعلومات.

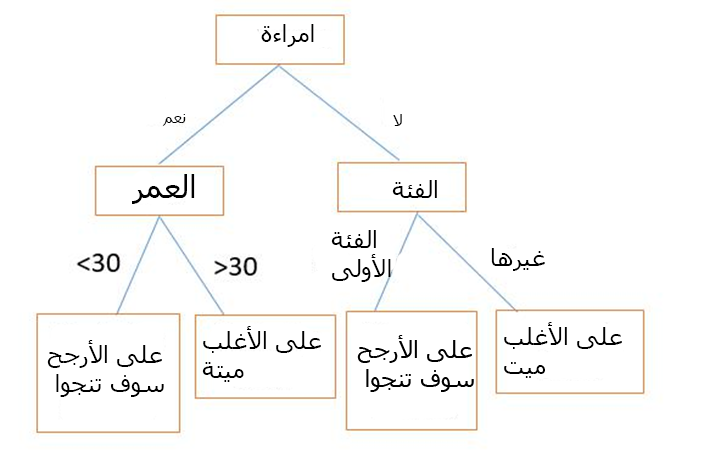
Let’s take a look at our Titanic variables:



Source: [Kaggle](https://www.kaggle.com/c/titanic/data)

نحن نبني شجرة من خلال اختيار أحد هذه المتغيرات وتقسيم مجموعة البيانات وفقًا لذلك.

Machine Learning for Humans



الأنقسام الأول يفصل مجموعة البيانات الخاصة بنا إلى رجال ونساء. ثم ، يتم تقسيم فرع النساء مرة أخرى في العمر (الانقسام الذي يقلل من الفوضى). وبالمثل ، يتم تقسيم فرع الرجال حسب الفصل. باتباع الشجرة لراكب جديد ، يمكنك استخدام الشجرة لتخمين ما إذا كانت قد توفيت أم لا.

مثال تيتانيك هو حل مشكلة التصنيف ("البقاء" أو "الموت"). إذا كنا نستخدم أشجار القرار من أجل الحسر - على سبيل المثال ، للتنبؤ بأسعار المساكن - فإننا سننشئ انقسامات حول أهم الميزات التي تحدد أسعار المساكن. كم قدم مربع: أكثر من أو أقل من \_\_\_؟ كم عدد غرف النوم والحمامات: أكثر من أو أقل من \_\_\_؟

بعد ذلك ، أثناء الاختبار ، ستقوم بختبار منزلًا محددًا من خلال جميع الانقسامات وتأخذ متوسط جميع أسعار المساكن في النقطة النهائية (العقدة السفلية) حيث ينتهي المنزل باعتباره تنبؤك لسعر البيع.

Machine Learning for Humans

نماذج أشجار القرار ، بما في ذلك max\_depth و max\_leaf\_nodes. راجع وحدة scikit-learn module التعلم المعرفية على أشجار القرار للحصول على نصائح حول تحديد هذه المعلمات

.

تعتبر أشجار القرار فعالة لأنها سهلة القراءة وقوية حتى مع البيانات الفوضوية ورخيصة من الناحية الحسابية للنشر مرة واحدة بعد التدريب. تعد أشجار القرار جيدة أيضًا للتعامل مع البيانات المختلطة (العددية أو الفئوية).

. ومع ذلك ، فإن أشجار القرارات مكلفة للتدريب من الناحية الحسابية ، وهي تنطوي على مخاطر كبيرة من الأكداء ، وتميل إلى العثور على القيمة صغرى المحلية لأنها لا تستطيع العودة بعد أن تنقسم. لمعالجة نقاط الضعف هذه ، ننتقل إلى طريقة توضح قوة دمج العديد من أشجار القرار في نموذج واحد.

غابة عشوائية: مجموعة من أشجار القرار

. يسمى النموذج الذي يتألف من العديد من النماذج نموذجًا للمجموعة ، وعادة ما تكون هذه إستراتيجية رابحة.

: يمكن لشجرة قرار واحدة القيام بالكثير من الأخطاء لأنها تحتوي على أحكام بالأبيض والأسود. الغابة العشوائية عبارة عن مقدر تقديري يقوم بتجميع العديد من أشجار القرارات ، مع بعض التعديلات المفيدة:

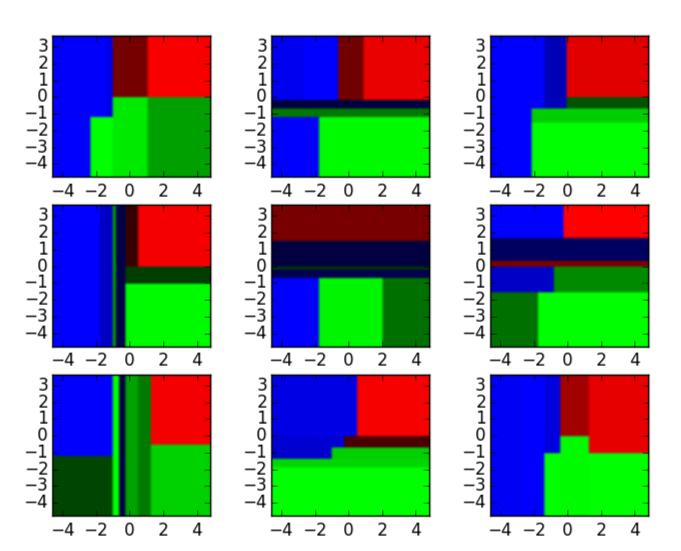
1. يقتصر عدد المميزات التي يمكن تقسيمها في كل عقدة على بعض النسبة المئوية من الإجمالي (هذا مقياس تشعبي يمكنك اختياره - راجع وثائق scikit-Learn للحصول على التفاصيل). هذا يضمن أن نموذج الفرقة لا يعتمد بشكل كبير على أي ميزة فردية ، ويستخدم بشكل عادل جميع الميزات المحتملة التنبؤية.

.2 ترسم كل شجرة عينة عشوائية من مجموعة البيانات الأصلية عند توليد انشقاقاتها ، مما يضيف عنصرًا إضافيًا من العشوائية يمنع الإفراط في التركيب.

تمنع هذه التعديلات أيضًا الترابط الشديد بين الأشجار. بدون رقم 1 ورقم 2 أعلاه ، ستكون كل شجرة متطابقة ، لأن الانقسام الثنائي التعاودي أمر حتمي.

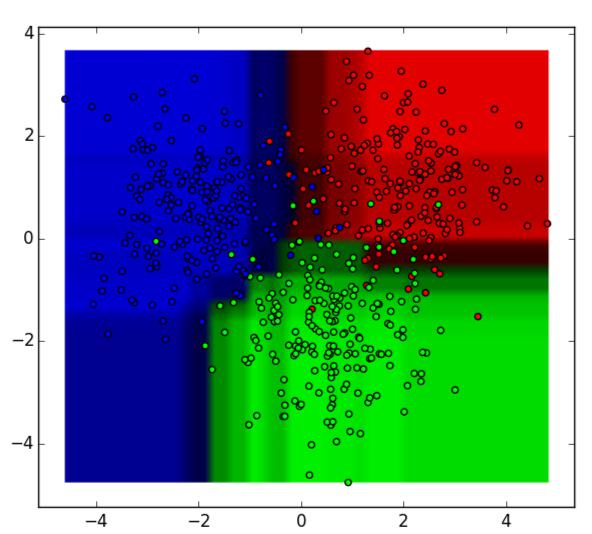
Machine Learning for Humans

. للتوضيح ، راجع تصنيفات أشجار القرار التسع هذه أدناه.



Source: [http://xenon.stanford.edu/~jianzh/ml/](Source:%20http://xenon.stanford.edu/~jianzh/ml/)

يمكن تجميع مصنفات شجرة القرار هذه في مجموعة غابة عشوائية تجمع بين مدخلاتها. فكر في المحاور الأفقية والرأسية لكل ناتج من شجرة القرار كميزات x1 و x2. عند قيم معينة لكل ميزة ، تقوم شجرة القرار بإخراج تصنيف "أزرق" ، "أخضر" ، "أحمر" ، إلخ.



Source: <http://xenon.stanford.edu/~jianzh/ml/>

Machine Learning for Humans

يتم تجميع هذه النتائج ، من خلال الأصوات النموذجية أو المتوسط ، في نموذج مجموعة واحد ينتهي به الأمر إلى التفوق في الأداء على مخرجات أي شجرة قرارات فردية.

تعد الغابات العشوائية نقطة انطلاق ممتازة لعملية النمذجة ، نظرًا لأنها تميل إلى الأداء القوي مع التسامح العالي للبيانات الأقل تنظيفًا ويمكن أن تكون مفيدة لمعرفة الميزات المهمة بالفعل من بين العديد من الميزات.

هناك العديد من نماذج المجموعات الذكية الأخرى التي تجمع بين أشجار القرار وتنتج أداءً ممتازًا - تحقق من XGBoost (Extreme Gradient Boosting) كمثال.

And with that, we conclude our study of supervised learning!

Nice work. In this section we’ve covered:

* Two **non-parametric** supervised learning algorithms: **k-NN** and **decision** **trees**
* Measures of **distance** and **information gain**
* **Random forests**, which are an example of an **ensemble model**
* **Cross-validation** and **hyperparameter tuning**

Hopefully, you now have some solid intuitions for how we learn *f* given a training data set and use this to make predictions with the test data.

Next, we’ll talk about how to approach problems where we don’t have any labeled training data to work with, in [Part 3: Unsupervised Learning](https://medium.com/@v_maini/unsupervised-learning-f45587588294).

Practice materials & further reading

**2.3a — Implementing k-NN**

Try this [walkthrough](http://machinelearningmastery.com/tutorial-to-implement-k-nearest-neighbors-in-python-from-scratch/) for implementing k-NN from scratch in Python. You may also want to take a look at the [scikit-learn](http://scikit-learn.org/stable/modules/generated/sklearn.neighbors.KNeighborsClassifier.html) documentation to get a sense of how pre-built implementations work.

Machine Learning for Humans

2.3b— أشجار القرار

جرب معمل أشجار القرار في الفصل 8 من مقدمة إلى التعلم الإحصائي. يمكنك أيضًا اللعب بمجموعة بيانات Titanic ، والتحقق من هذا البرنامج التعليمي الذي يغطي نفس المفاهيم المذكورة أعلاه مع الكود المصاحب. هنا هو تنفيذ scikit-Learn للغابة العشوائية للاستخدام خارج الصندوق في مجموعات البيانات

الفصل ثالث : تعلم الغير مشرف



*)* *التعلم غير مشرف تحليل عنقودي: خوارزمية تصنيفية ، تجميع هرمي. تخفيض الأبعاد : تحليل العناصر الرئيسي (تار) , مجزئ القيمة المفردة (جيم).*

كيف تجد البنية الأساسية لمجموعة البيانات؟ كيف تلخصها وتضعها في مجموعات أكثر فائدة؟ كيف يمكنك تمثيل البيانات بشكل فعال في صيغة مضغوط؟ هذه هي أهداف التعلم غير مشرف ، والذي يسمى "بدون إشراف" لأنك تبدأ ببيانات غير مصنفة (لا يوجد ق)

لا ي

تتمثل مهمتا التعلم غير مشرف التان سنستكشفهما في تجميع البيانات في مجموعات عن طريق التشابه وتقليل الأبعاد لضغط البيانات مع الحفاظ على هيكلها وفائدتها.

أمثلة على الأماكن التي قد تكون فيها طرق التعلم غير مشرف مفيدة:

* تقسم منصة إعلانية سكان الولايات المتحدة إلى مجموعات أصغر لها نفس الخصائص الديمغرافية وعادات الشراء بحيث يمكن للمعلنين الوصول إلى السوق المستهدفة من خلال الإعلانات ذات الصلة.
* تقوم Airbnb بتجميع قوائم الإسكان الخاصة بها في الأحياء بحيث يمكن للمستخدمين التنقل في القوائم بسهولة أكبر.
* يقلل فريق علم البيانات من عدد الأبعاد في مجموعة بيانات كبيرة لتبسيط النمذجة وتقليل حجم الملف.

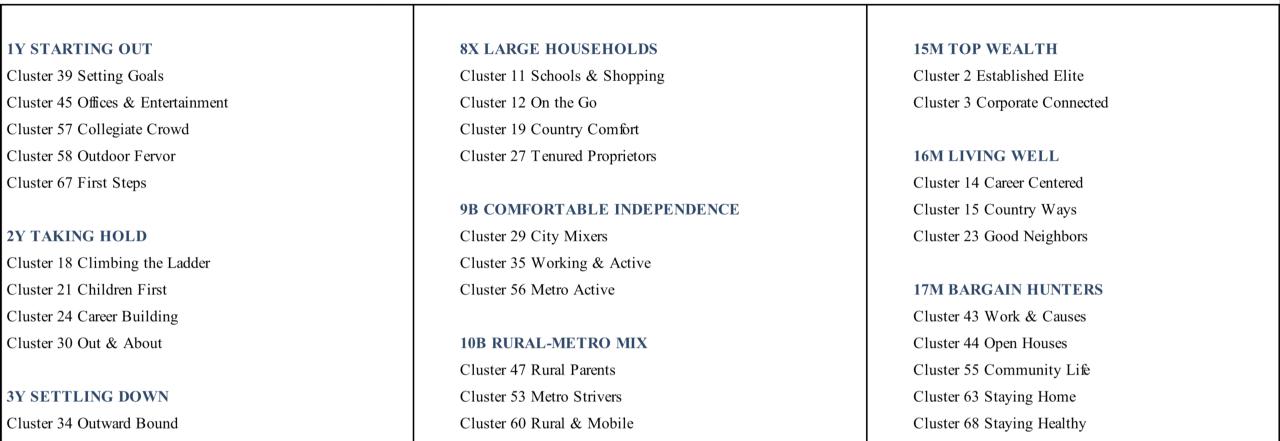
55

Machine Learning for Humans

على عكس التعلم المشرف ، ليس من السهل دائمًا وضع مقاييس لمدى جودة أداء خوارزمية التعلم غير الخاضعة للإشراف. غالبًا ما يكون "الأداء" شخصيًا ومخصصًا للمجال.

التحليل العنقودي

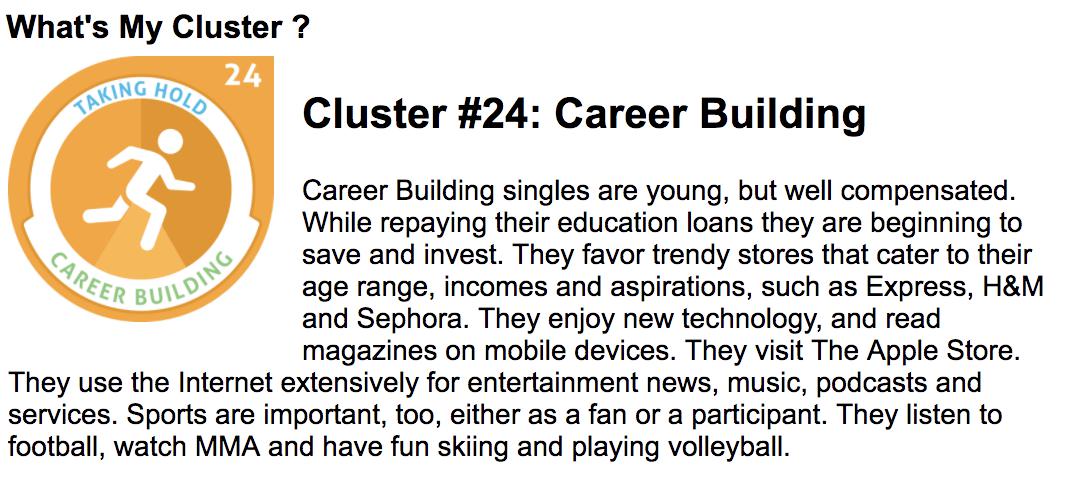
أحد الأمثلة المثيرة للاهتمام على التجميع في العالم الحقيقي هو نظام تجميع مراحل الحياة لمزود البيانات Acxiom ، Personicx. تقسم هذه الخدمة الأسر الأمريكية إلى 70 مجموعة متميزة ضمن 21 مجموعة مراحل الحياة التي يستخدمها المعلنون عند استهداف إعلانات Facebook ، والإعلانات المصورة ، وحملات البريد المباشر ، إلخ.



A selection of Personicx demographic clusters

يكشف المستند التعريفي التمهيدي أنهم استخدموا تجميع النقط الوسطى وتحليل المكونات الرئيسية ، وكلاهما من التقنيات التي تم تناولها في هذا القسم.

يمكنك أن تتخيل كيف أن الوصول إلى هذه المجموعات مفيد للغاية للمعلنين الذين يرغبون في (1) فهم قاعدة عملائهم الحاليين و (2) استخدام إنفاقهم الإعلاني بشكل فعال من خلال استهداف عملاء جدد محتملين مع الخصائص الديمغرافية والاهتمامات وأنماط الحياة ذات الصلة.



You can actually find out which cluster you personally would belong to by answering a few simple questions in Acxiom’s

[“What’s My Cluster?” tool](https://isapps.acxiom.com/personicx/personicx.aspx).

Let’s walk through a couple of clustering methods to develop intuition for how this task can be performed.

k-means clustering

*“And k rings were given to the race of Centroids, who above all else, desire power.”*

The goal of clustering is to create groups of data points such that points in different clusters are dissimilar while points within a cluster are similar.

With **k-means clustering**, we want to cluster our data points into k groups. A larger k creates smaller groups with more granularity, a lower k means larger groups and less granularity.

The output of the algorithm would be a set of “labels” assigning each data point to one of the k groups. In k-means clustering, the way these groups are defined is by creating a **centroid** for each group. The centroids are like the heart of the cluster, they “capture” the points closest to them and add them to the cluster.

Machine Learning for Humans

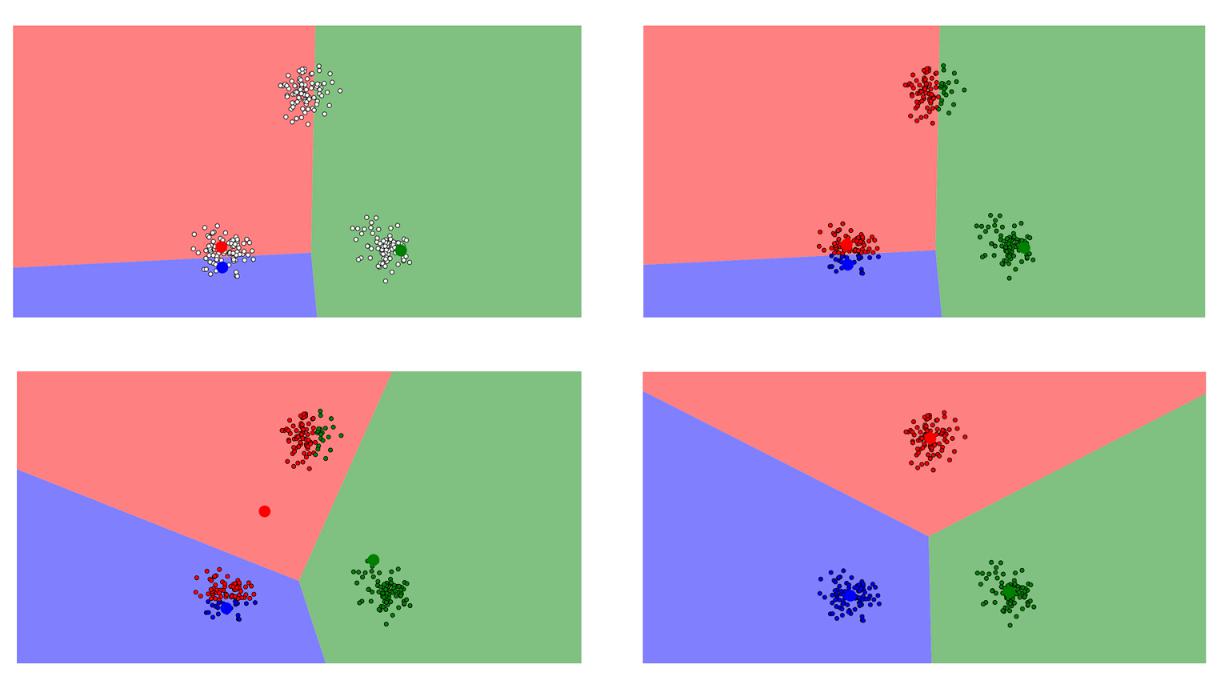
Think of these as the people who show up at a party and soon become the centers of attention because they’re so magnetic. If there’s just one of them, everyone will gather around; if there are lots, many smaller centers of activity will form.

Here are the steps to k-means clustering:

1. Define the k centroids. Initialize these at random (there are also fancier algorithms for initializing the centroids that end up converging more effectively).
2. Find the closest centroid & update cluster assignments. Assign each data point to one of the k clusters. Each data point is assigned to the nearest centroid’s cluster. Here, the measure of “nearness” is a hyperparameter — often Euclidean distance.
3. Move the centroids to the center of their clusters. The new position of each centroid is calculated as the average position of all the points in its cluster.

Keep repeating steps 2 and 3 until the centroid stop moving a lot at each iteration (i.e., until the algorithm converges)

That, in short, is how k-means clustering works! Check out this [visualization](https://www.naftaliharris.com/blog/visualizing-k-means-clustering/) of the algorithm — read it like a comic book. Each point in the plane is colored according the centroid that it is closest to at each moment. You’ll notice that the centroids (the larger blue, red, and green circles) start randomly and then quickly adjust to capture their respective clusters.



Another real-life application of k-means clustering is classifying handwritten digits. Suppose we have images of the digits as a long vector of pixel brightnesses. Let’s say the images are black and white and are 64x64 pixels. Each pixel represents a dimension. So the world these images live in has 64x64=4,096 dimensions. In this 4,096-dimensional world, k-means clustering allows us to group the images that are close together and assume they represent the same digit, which can achieve [pretty](http://ieeexplore.ieee.org/document/6755106/?reload=true) [good results](http://ieeexplore.ieee.org/document/6755106/?reload=true) for digit recognition.

Hierarchical clustering

*“Let’s make a million options become seven options. Or five. Or twenty? Meh, we can decide later.”*

Hierarchical clustering is similar to regular clustering, except that you’re aiming to build a hierarchy of clusters. This can be useful when you want flexibility in how many clusters you ultimately want. For example, imagine grouping items on an online marketplace like Etsy or Amazon. On the homepage you’d want a few broad categories of items for

simple navigation, but as you go into more specific shopping categories you’d want increasing levels of granularity, i.e. more distinct clusters of items.

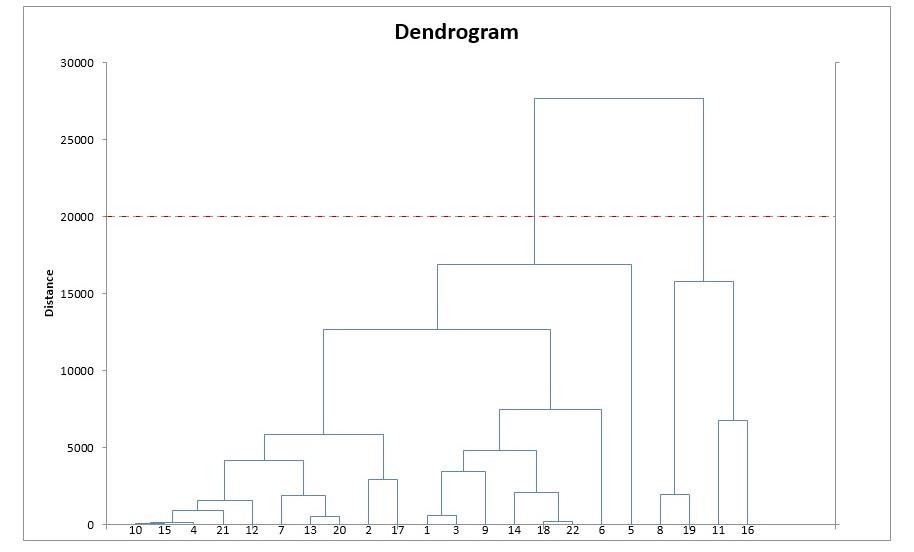
In terms of outputs from the algorithm, in addition to cluster assignments you also build a nice tree that tells you about the hierarchies between the clusters. You can then pick the number of clusters you want from this tree.

Here are the steps for hierarchical clustering:

1. **Start with N clusters,** one for each data point.
2. **Merge the two clusters that are closest to each other.** Now you have N-1 clusters.
3. **Recompute the distances between the clusters.** There are several ways to dothis (see this tutorial for more details). One of them (called average-linkage clustering) is to consider the distance between two clusters to be the average distance between all their respective members.
4. **Repeat steps 2 and 3 until you get one cluster of N data points.** You get a tree(also known as a dendrogram) like the one below.
5. **Pick a number of clusters and draw a horizontal line in the dendrogram.**

For example, if you want k=2 clusters, you should draw a horizontal line around “distance=20000.” You’ll get one cluster with data points 8, 9, 11, 16 and one cluster with the rest of the data points. In general, the number of clusters you get is the number of intersection points of your horizontal line with the vertical lines in the dendrogram.

Machine Learning for Humans



Source: [Solver.com](https://www.solver.com/hierarchical-clustering-example). For more detail on hierarchical clustering, you can check [this video](https://www.youtube.com/watch?v=OcoE7JlbXvY) out.

Dimensionality reduction

*“It is not the daily increase, but the daily decrease. Hack away at the unessential.”—Bruce Lee*

Dimensionality reduction looks a lot like compression. This is about trying to reduce the complexity of the data while keeping as much of the relevant structure as possible. If you take a simple 128 x 128 x 3 pixels image (length x width x RGB value), that’s 49,152 dimensions of data. If you’re able to reduce the dimensionality of the space in which these images live without destroying too much of the meaningful content in the images, then you’ve done a good job at dimensionality reduction.

We’ll take a look at two common techniques in practice: **principal component analysis**

and **singular value decomposition**.



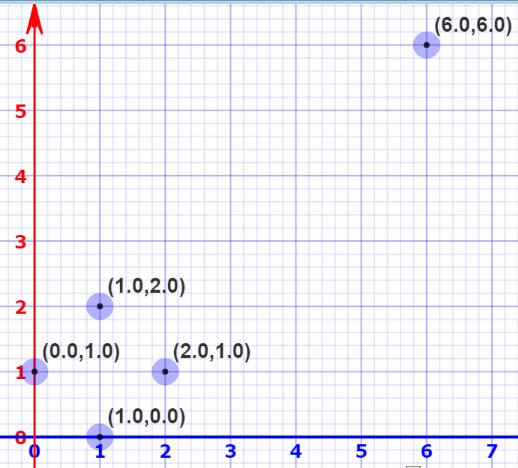
61

Part 3: Unsupervised Learning

Principal component analysis (PCA)

First, a little linear algebra refresher — let’s talk about **spaces** and **bases**.

You’re familiar with the coordinate plane with origin O(0,0) and **basis vectors** i(1,0) and j(0,1). It turns out you can choose a completely different basis and still have all the math work out. For example, you can keep O as the origin and choose the basis to vectors i’=(2,1) and j’=(1,2). If you have the patience for it, you’ll convince yourself that the point labeled (2,2) in the i’, j’ coordinate system is labeled (6, 6) in the i, j system.



Plotted using Mathisfun’s “[Interactive Cartesian Coordinates](https://www.mathsisfun.com/data/cartesian-coordinates-interactive.html)”

This means we can change the basis of a space. Now imagine much higher-dimensional space. Like, 50K dimensions. You can select a basis for that space, and then select only the 200 most significant vectors of that basis. These basis vectors are called **principal components**, and the subset you select constitute a new space that is smaller in dimensionality than the original space but maintains as much of the complexity of the data as possible.

To select the most significant principal components, we look at how much of the data’s variance they capture and order them by that metric.

Another way of thinking about this is that PCA remaps the space in which our data exists to make it more compressible. The transformed dimension is smaller than the original dimension.

By making use of the first several dimensions of the remapped space only, we can start gaining an understanding of the dataset’s organization. This is the promise of dimensionality reduction: reduce complexity (dimensionality in this case) while maintaining structure (variance). Here’s a [fun paper](http://bit.ly/2v0nxrA) Samer wrote on using PCA (and diffusion mapping, another technique) to try to make sense of the Wikileaks cable release.

Singular value decomposition (SVD)

Let’s represent our data like a big A = m x n matrix. SVD is a computation that allows us to decompose that big matrix into a product of 3 smaller matrices (U=m x r, diagonal matrix Σ=r x r, and V=r x n where r is a small number).

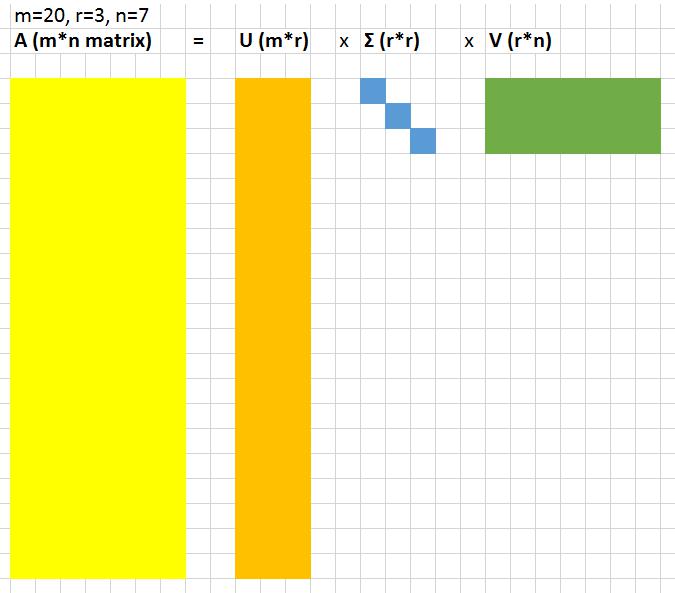


63

Part 3: Unsupervised Learning

Machine Learning for Humans

Here’s a more visual illustration of that product to start with:

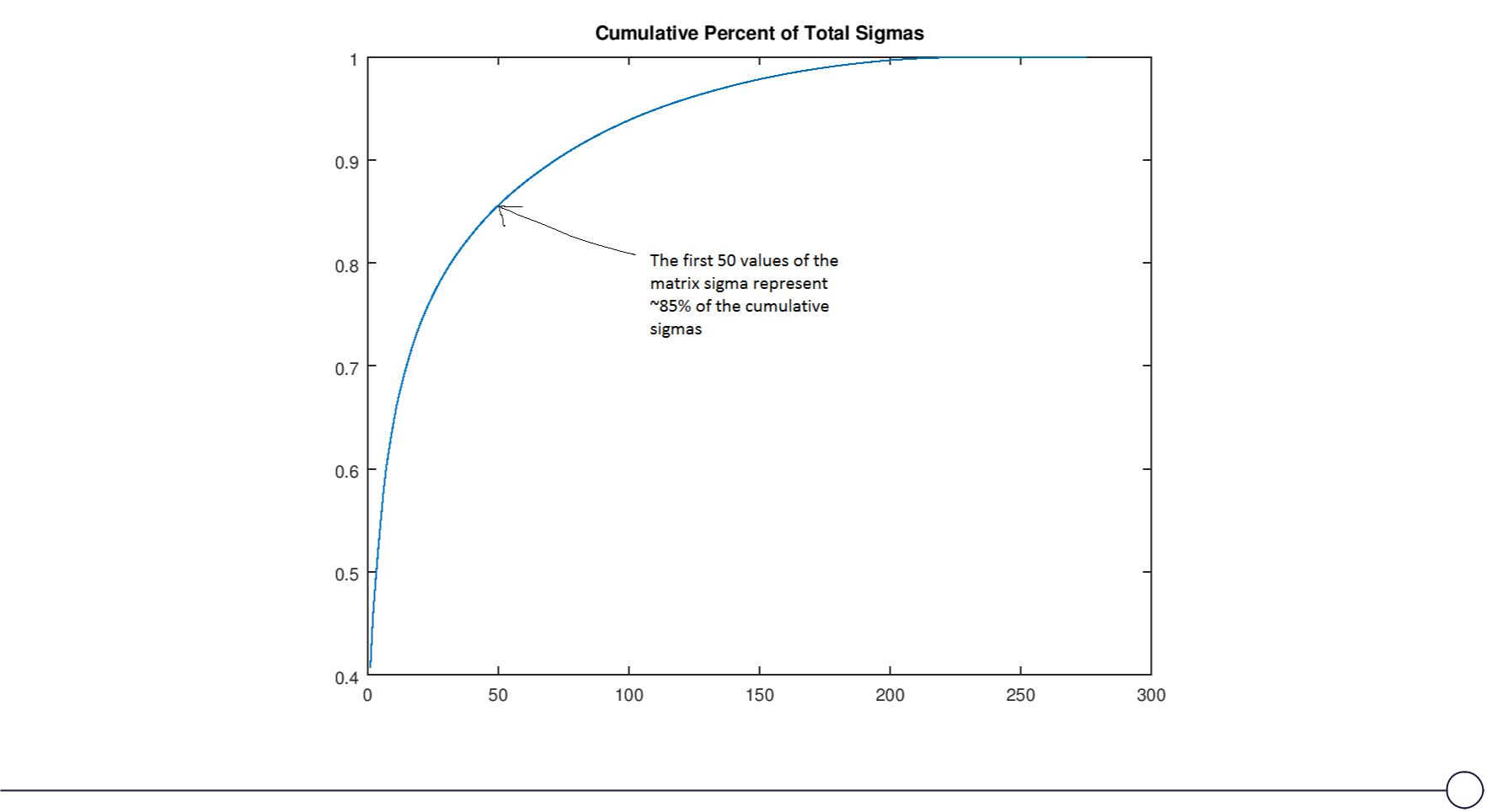


The values in the r\*r diagonal matrix Σ are called singular values. What’s cool about them is that these singular values can be used to compress the original matrix. If you drop the smallest 20% of singular values and the associated columns in matrices U and V, you save quite a bit of space and still get a decent representation of the underlying matrix.

To examine what that means more precisely, let’s work with this image of a dog:



We’ll use the code written in Andrew Gibiansky’s post on SVD. First, we show that if we rank the singular values (the values of the matrix Σ) by magnitude, the first 50 singular values contain 85% of the magnitude of the whole matrix Σ.

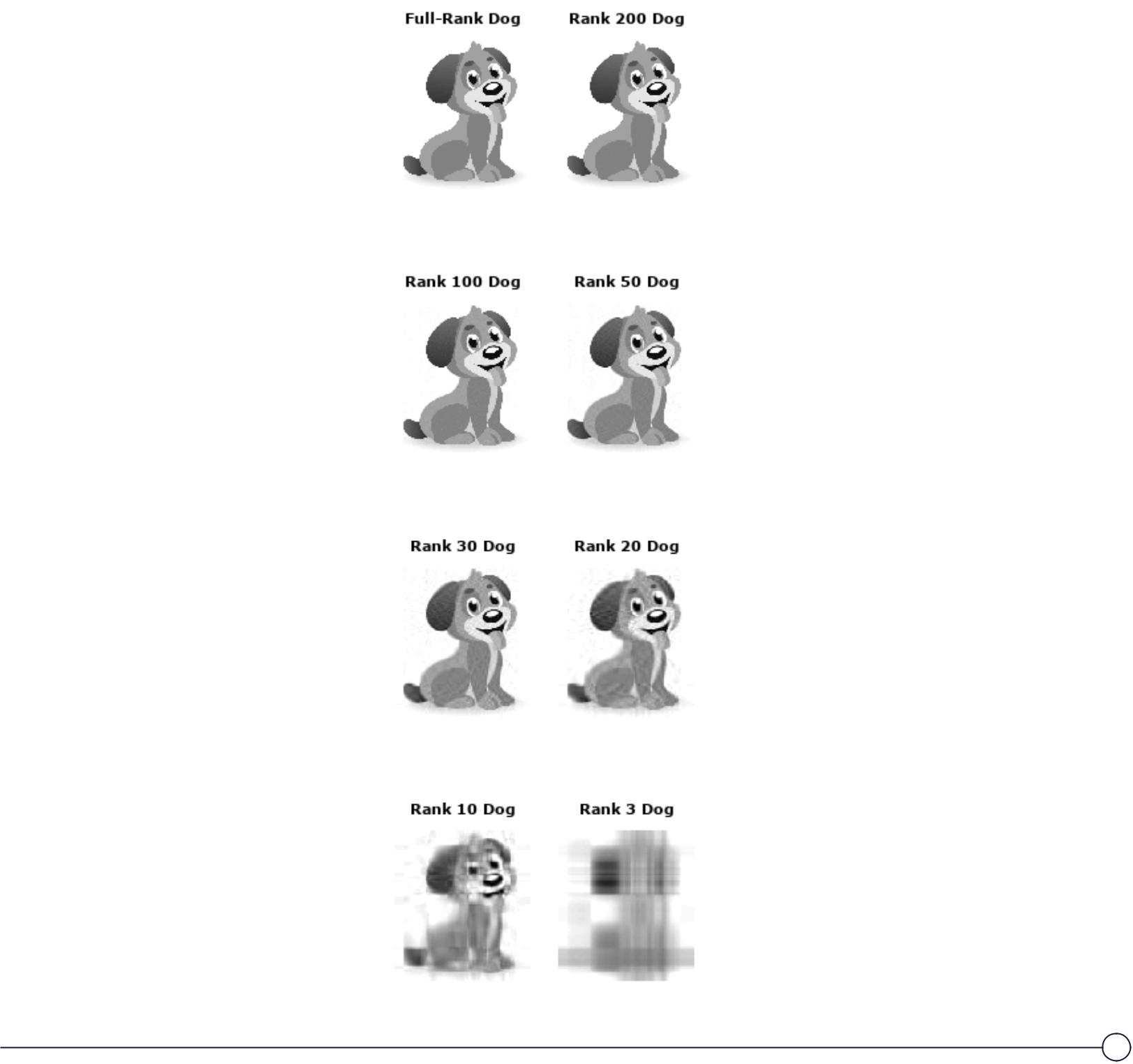


65

Part 3: Unsupervised Learning

Machine Learning for Humans

We can use this fact to discard the next 250 values of sigma (i.e., set them to 0) and just keep a “rank 50” version of the image of the dog. Here, we create a rank 200, 100, 50, 30, 20, 10, and 3 dog. Obviously, the picture is smaller, but let’s agree that the rank 30 dog is still good. Now let’s see how much compression we achieve with this dog. The original image matrix is 305\*275 = 83,875 values. The rank 30 dog is 305\*30+30+30\*275=17,430 — almost 5 times fewer values with very little loss in image quality. The reason for the calculation above is that we also discard the parts of the matrix U and V that get multiplied by zeros when the operation UΣ’V is carried out (where Σ’ is the modified version of Σ that only has the first 30 values in it).



66

Part 3: Unsupervised Learning

Machine Learning for Humans

Unsupervised learning is often used to preprocess the data. Usually, that means compressing it in some meaning-preserving way like with PCA or SVD before feeding it to a deep neural net or another supervised learning algorithm.

Onwards!

Now that you’ve finished this section, you’ve earned an awful, horrible, never-to-be-mentioned-again joke about unsupervised learning. Here goes…

*Person-in-joke-#1: Y would u ever need to use unsupervised tho?*

*Person-in-joke-#2: Y? there’s no Y.*

Next up… [Part 4: Neural Networks & Deep Learning!](https://medium.com/@v_maini/neural-networks-deep-learning-cdad8aeae49b)

Practice materials & further reading

3a—k-means clustering

Play around with this clustering [visualization](https://www.naftaliharris.com/blog/visualizing-k-means-clustering/) to build intuition for how the algorithm works. Then, take a look at this implementation of [k-means clustering](https://github.com/Datamine/MNIST-K-Means-Clustering) [for handwritten digits](https://github.com/Datamine/MNIST-K-Means-Clustering) and the associated tutorial.

3b—SVD

For a good reference on SVD, go no further than Andrew Gibiansky’s [post](http://andrew.gibiansky.com/blog/mathematics/cool-linear-algebra-singular-value-decomposition/).

Machine Learning for Humans

الجزء رابع : شبكات العصبية و تعلم العميق



*أين ، لماذا ، أين ، وكيف تعمل الشبكات العصبية العميقة. وحي من الدماغ. الشبكات العصبية التلافيفية (CNNs) والشبكات العصبية المتكررة (RNNs). تطبيقات العالم الحقيقي*

من خلال التعلم العميق ، ما زلنا نتعلم وظيفة f لتعيين الإدخال X لإخراج Y بأقل خسارة في بيانات الاختبار ، تمامًا كما كنا نفعل طوال الوقت. استرجع "بيان المشكلة" الأولي الخاص بنا من الجزء 2.1 حول التعلم الخاضع للإشراف:

**Y = f(X) +** ϵ

**Training:** machine learns*f*from labeled training data

**Testing:** machine predicts Y from unlabeled testing data

العالم الحقيقي هو فوضوي ، لذلك في بعض الأحيان و معقدة. في مشاكل اللغة الطبيعية ، تعني أحجام المفردات الكبيرة الكثير من الميزات. تتضمن مشكلات الرؤية الكثير من المعلومات المرئية حول البكسل. تتطلب ممارسة الألعاب اتخاذ قرار بناءً على سيناريوهات معقدة مع العديد من العقود المستقبلية المحتملة. تقنيات التعلم التي قمنا بتغطيتها جيدة حتى الآن عندما تكون البيانات التي نعمل بها غير معقدة بجنون ، ولكن ليس من الواضح كيف تم تعميمها على سيناريوهات مثل هذه.

التعلم العميق جيد حقًا في التعلم (ص) ، خاصةً في الحالات التي تكون فيها البيانات معقدة. في الواقع ، تُعرف الشبكات العصبية المصطنعة بأنها مُقَرِّبة الأقترانات العامة لأنها قادرة على [تعلم أي اقتران](http://neuralnetworksanddeeplearning.com/chap4.html) ، بصرف النظر عن مدى شكلها ، **باستخدام طبقة واحدة مخفية فقط.**

دعونا نلقي نظرة على مشكلة تصنيف الصور. نلتقط صورة كمدخلات ، ونخرج فئة (على سبيل المثال ، كلب ، قطة ، سيارة).

Machine Learning for Humans

بشكل صوري شبكة عصبية عميقة لتصنيف صور تبدو بهذا شكل:

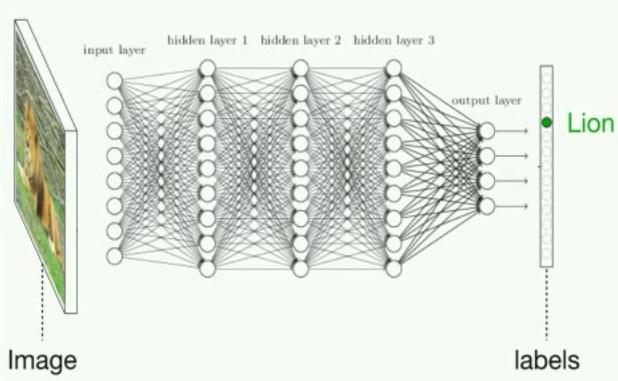


Image from Jeff Clune’s 1-hour [Deep Learning Overview](https://www.youtube.com/watch?v=3lp9eN5JE2A&t=1631s) on YouTube

لكن في الحقيقة ، هذه معادلة رياضية ضخمة بها ملايين المصطلحات والكثير من المعلمات. المدخل **س** هو ، على سبيل المثال ، صورة ذات تدرجات رمادية ممثلة بمصفوفة و\*ه من سطوع البكسل. الإخراج **ص** هو متجه من للاحتمالات فئة. هذا يعني أننا كمخرج احتمالية أن يكون كل فصل هو التسمية الصحيحة. إذا كانت هذه الشبكة العصبية تعمل بشكل جيد ، فيجب أن يكون أكبر احتمال للفئة الصحيحة. وتقوم الطبقات الموجودة في الوسط بمجموعة من الضرب المصفوفي عن طريق جمع عمليات التنشيط x الأوزان مع التحولات غير الخطية (اقتران تفعيل) بعد كل طبقة مخفية لتمكين الشبكة من تعلم وظيفة غير خطية.

بشكل لا يصدق ، يمكنك استخدام انحدار التدرج بنفس الطريقة التي فعلنا بها مع الحسر الخطي في الجزء 2.1 لتدريب هذه المعلمات بطريقة تقلل من الخسارة. لذلك مع الكثير من الأمثلة والكثير من النسب التدرج ، يمكن أن يتعلم النموذج كيفية تصنيف صور الحيوانات بشكل صحيح. وهذا باختصار باختصار هو "التعلم العميق".

Machine Learning for Humans

لماذا اداء تعلم العميق جيد و بعض من تاريخه

لقد كانت الشبكات العصبية الاصطناعية موجودة بالفعل لفترة طويلة. يشار تاريخيا إلى تطبيقها على أنها **علم التحكم الآلي** (1940s - 1960s) ، **والربط** (الثمانينيات والتسعينيات) ، ثم أصبحت رواجًا كحلقة تعليمية عميقة 2006 عندما بدأت الشبكات العصبية تتطور ، "[أعمق](http://www.deeplearningbook.org/contents/intro.html)" لكن في الآونة الأخيرة فقط بدأنا بالفعل في خدش سطح إمكاناتها الكاملة.

كما وصف [Andrej Karpathy](https://cs.stanford.edu/people/karpathy/) (مدير الذكاء الاصطناعى في تسلا ، الذي نميل إلى التفكير به باعتباره ساحر للتعلم العميق) ، هناك بشكل عام "أربعة عوامل منفصلة تعيق الذكاء الاصطناعي:"

1. الحوسبة (قانون مور)
2. البيانات (بشكل مناسب و مرتب)
3. الخورزميات
4. البنية تحتية

!) في العقد الماضي أو نحو ذلك ، تم أخيرًا اطلاق الإمكانات الكاملة للتعلم العميق من خلال التقدم في (1) و (2) ، مما أدى بدوره إلى مزيد من تقدم في (3) و(4) - وهكذا تستمر الدورة ، حيث يتجمع عدد أكبر من البشر بشكل متصاعد في الخطوط الأمامية لأبحاث التعلم العميق على طول الطريق (فقط فكر فيما تفعله الآن!)

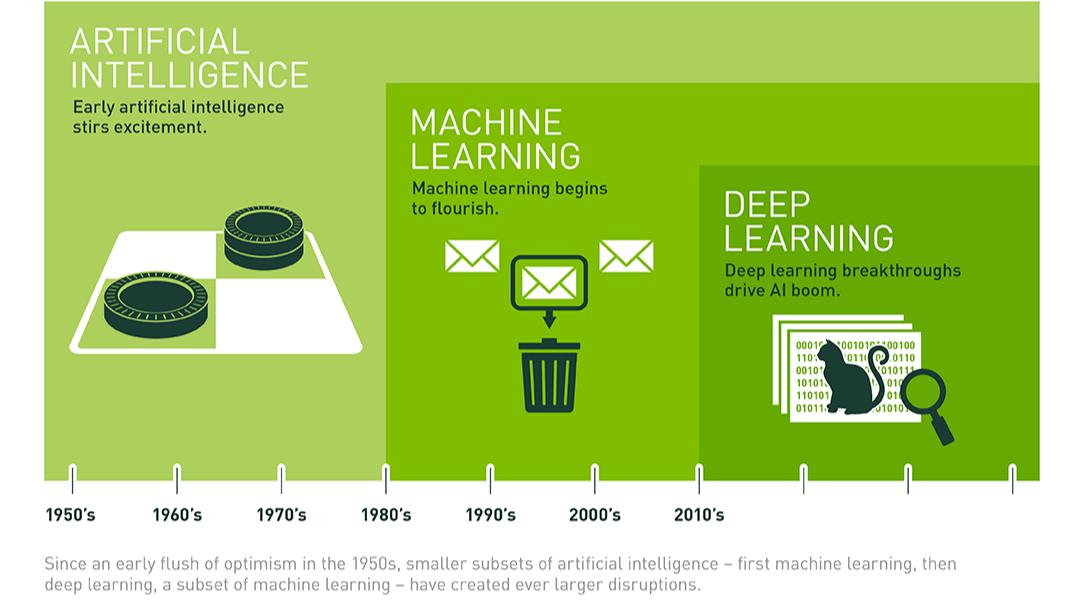


Illustration by [NVIDIA](https://blogs.nvidia.com/blog/2016/07/29/whats-difference-artificial-intelligence-machine-learning-deep-learning-ai/), a leading maker of graphics processing units (GPUs) which were originally built for for gaming but turned out to be well-suited to the type of [parallel computing](http://www.nvidia.com/object/what-is-gpu-computing.html) required by deep neural networks

في بقية هذا القسم ، سنقدم بعض المعلومات الأساسية عن البيولوجيا والأحصاء لشرح ما يحدث داخل الشبكات العصبية ، ومن ثم التحدث عبر بعض التطبيقات المدهشة للتعلم العميق. أخيرًا ، سنربط بعض الموارد حتى تتمكن من تطبيق التعلم العميق بنفسك ، حتى الجلوس على الأريكة في البيجامة مع جهاز كمبيوتر محمول ، من أجل تحقيق أداء أعلى من المستوى البشري بسرعة على أنواع معينة من المشاكل.

الحصول على إلهام من الدماغ (أم أنه مجرد إحصائيات؟) - ماذا يحدث داخل الشبكات العصبية

العصبونات , تعلم المميزات , وطبقات من تجريد

بينما تقرأ هذه الكلمات ، فأنت لا تفحص كل حرف من كل كلمة ، أو كل بكسل يكون كل حرف ، لكي تعرف معنى الكلمة. أنت تجرد

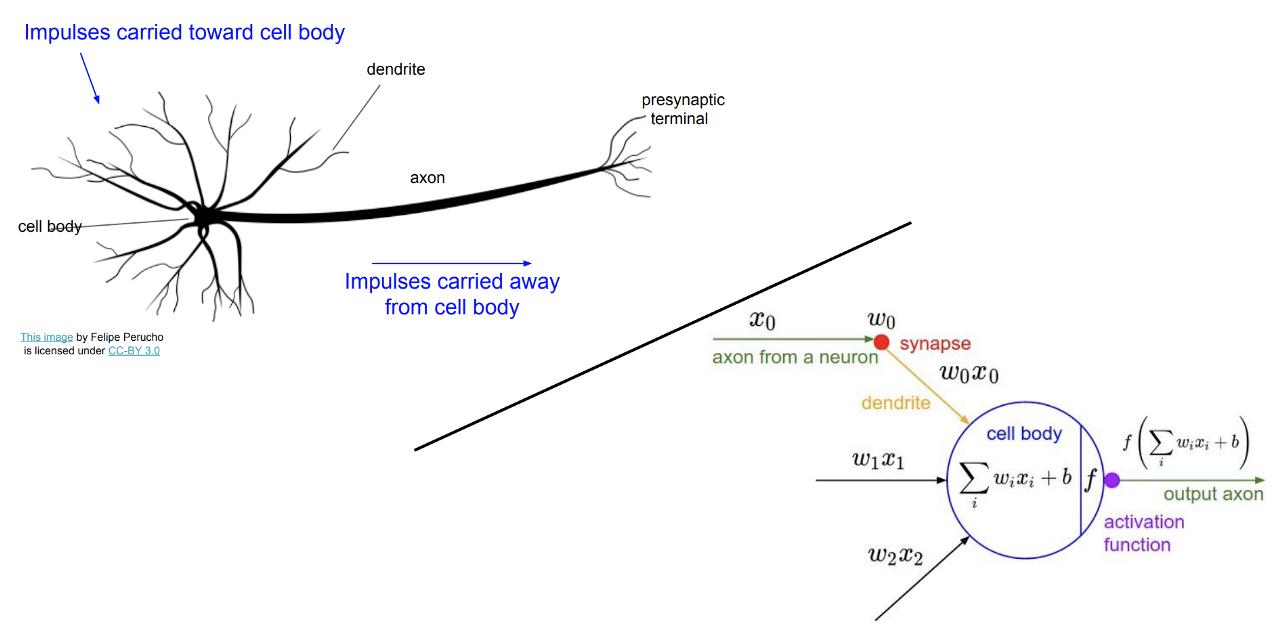
Machine Learning for Humans

بعيدًا عن التفاصيل وتجميع الأشياء في مفاهيم عليا المستوى: الكلمات والعبارات والجمل والفقرات.

Tihs abiilty to exaimne hgiher-lveel fteaures is waht aollws yuo to unedrtsand waht is hpapening in tihs snetecne wthiout too mcuh troulbe (or myabe yuo sned too mnay dnruk txets).

يحدث الشيء نفسه في الرؤية ، ليس فقط في البشر ولكن في النظم البصرية للحيوانات بشكل عام.

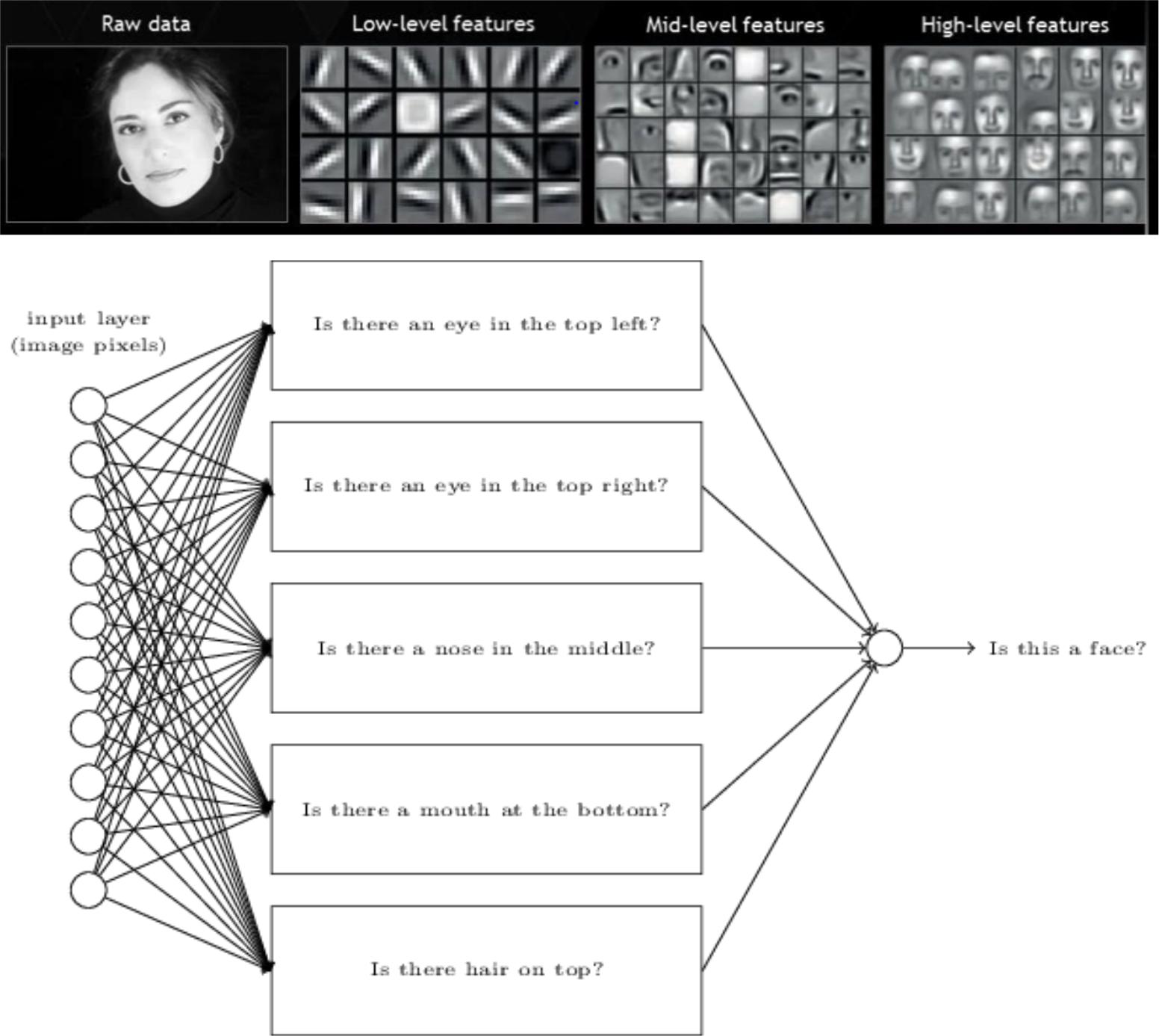
تتكون الأدمغة من الخلايا العصبية التي "تطلق" عن طريق إرسال إشارات كهربائية إلى الخلايا العصبية الأخرى بعد "تنشيطها" بشكل كافٍ. هذه الخلايا العصبية قابلة للطرق من حيث مقدار الإشارة من الخلايا العصبية الأخرى التي ستضيف إلى مستوى تنشيط الخلايا العصبية (بمعنى غامض ، ينتهي الأمر بتدريب الأوزان التي تربط الخلايا العصبية ببعضها البعض لجعل الروابط العصبية أكثر فائدة ، تمامًا مثل المعلمات في الانحدار الخطي يمكن تدريبه لتحسين التعيين من المدخلات إلى المخرجات).



Side-by-side illustrations of biological and artificial neurons, via [Stanford’s CS231n](http://cs231n.stanford.edu/slides/2017/cs231n_2017_lecture4.pdf). This analogy can’t be taken too literally — biological neurons can do things that artificial neurons can’t, and vice versa — but it’s useful to understand the biological inspiration. See Wikipedia’s description of [biological vs. artificial neurons](https://en.wikipedia.org/wiki/Artificial_neuron#Biological_models) for more detail.

Machine Learning for Humans

Our biological networks are arranged in a hierarchical manner, so that certain neurons end up detecting not extremely specific features of the world around us, but rather more abstract features, i.e. patterns or groupings of more low-level features. For example, the [fusiform face area](https://en.wikipedia.org/wiki/Fusiform_face_area) in the human visual system is specialized for facial recognition.



Top: Illustration of learning increasingly abstract features, via [NVIDIA](https://www.slideshare.net/NVIDIA/gpuaccelerated-deep-learning-for-cudnn-v2). Bottom: diagram of how an artificial neural network takes raw pixel inputs, develops intermediate “neurons” to detect higher-level features (e.g. presence of a nose), and combines the outputs of these to create a final output. Illustration from Neural Networks and Deep Learning

([Nielsen, 2017](http://neuralnetworksanddeeplearning.com/chap1.html)).



73

Part 4: Neural Networks & Deep Learning

Machine Learning for Humans

تم اكتشاف هذا الهيكل الهرمي الذي عرضته الشبكات العصبية البيولوجية في الخمسينيات من القرن الماضي عندما كان الباحثان David Hubel و Torsten Wiesel يدرسان الخلايا العصبية في القشرة البصرية للقطط. لم يتمكنوا من ملاحظة التنشيط العصبي بعد تعريض القطة لمجموعة متنوعة من المحفزات: البقع الداكنة ، والبقع المضيئة ، والتلويح باليد ، وحتى صور النساء في المجلات. لكن في إحباطهم ، عندما أزالوا شريحة من جهاز العرض بزاوية قطرية ، لاحظوا بعض النشاط العصبي! اتضح أن الحواف المائلة بزاوية معينة تتسبب في تنشيط بعض الخلايا العصبية.



Background via [Knowing Neurons](http://knowingneurons.com/2014/10/29/hubel-and-wiesel-the-neural-basis-of-visual-perception/)

هذا منطقي من الناحية التطورية لأن البيئات الطبيعية بشكل عام صاخبة وعشوائية (تخيل سهلًا عشبيًا أو تضاريس صخرية). لذلك عندما يرى قط في البرية "حافة" ، أي خط يتناقض مع خلفيته ، فقد يشير ذلك إلى وجود كائن أو مخلوق في المجال البصري. عندما يتم تنشيط توليفة معينة من عصبونات الحواف ، فإن تلك التنشيطات سوف تتحد لتنتج تنشيطًا أكثر تجريدًا ، وهكذا دواليك ، حتى يصبح التجريد النهائي مفهومًا مفيدًا ، مثل "طائر" أو "ذئب".

الفكرة وراء الشبكة العصبية العميقة هي محاكاة بنية مماثلة بطبقات من الخلايا العصبية الاصطناعية.

لماذا لا تعمل النماذج الخطية

To draw from Stanford’s excellent deep learning course, [CS231n: Convolutional Neural](http://cs231n.stanford.edu/syllabus.html) [Networks and Visual Recognition](http://cs231n.stanford.edu/syllabus.html), imagine that we want to train a neural network to classify images with the correct one of the following labels: *["plane", "car", "bird",* *"cat", "deer", "dog", "frog", "horse", "ship", "truck"].*

One approach could be to construct a “template”, or average image, of each class of image using the training examples, and then use a nearest-neighbors algorithm during

testing to measure the distance of each unclassified image’s pixel values, in aggregate, to each template. This approach involves no layers of abstraction. It’s a linear model that combines all the different orientations of each type of image into one averaged blur.

For instance, it would take all the cars — regardless of whether they’re facing left, right, center, and regardless of their color — and average them. The template then ends up looking rather vague and blurry.



Example drawn from Stanford’s [CS231n: Convolutional Neural Networks and Visual Recognition](http://cs231n.stanford.edu/syllabus.html),

Lecture 2.

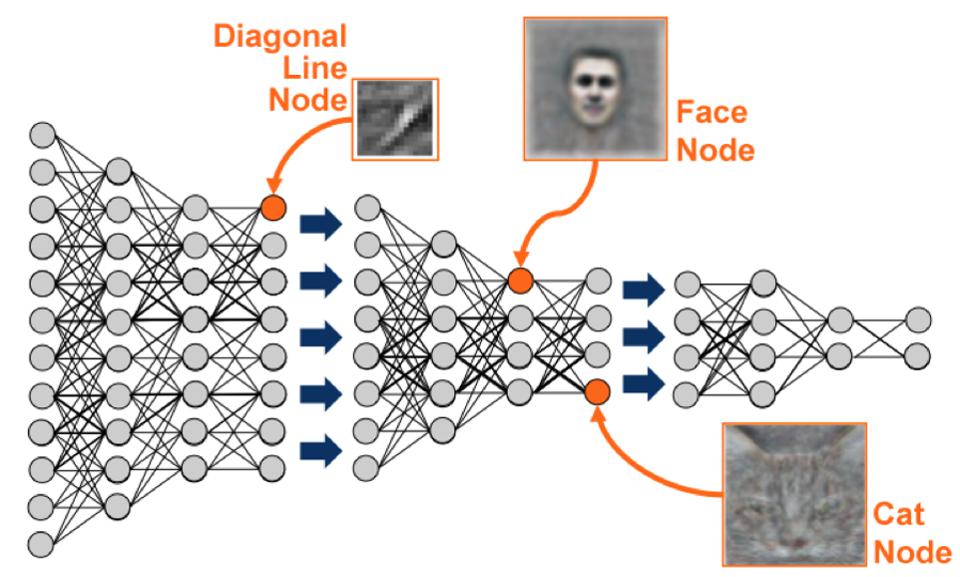
Notice that the horse template above appears to have two heads. This doesn’t really help us: we want to be able to detect right-facing horse or a left-facing horse separately, and then if either one of those features is detected, then we want to say we’re looking at a horse. This flexibility is provided by deep neural nets, as we will see in the next section.

Deep neural networks approach the image classification problem using layers of abstraction

To repeat what we explained earlier in this section: the input layer will take raw pixel brightnesses of an image. The final layer will be an output vector of **class probabilities** (i.e. the probability of the image being a “cat”, “car”, “horse”, etc.)

But instead of learning a simple linear model relating input to output, we’ll instead construct intermediate **hidden layers** of the network will learn increasingly abstract features, which enables us to not lose all the nuance in the complex data.

Machine Learning for Humans



Source: [Analytics Vidhya](https://www.analyticsvidhya.com/blog/2017/04/comparison-between-deep-learning-machine-learning/)

تمامًا كما وصفنا أدمغة الحيوانات التي تكتشف الميزات المجردة ، ستتعلم الخلايا العصبية الاصطناعية في الطبقات المخفية اكتشاف المفاهيم المجردة - أيًا كانت المفاهيم مفيدة في نهاية المطاف لالتقاط معظم المعلومات وتقليل مقدار في عدم دقة مخرج الشبكة (هذا مثال على ذلك) التعلم دون إشراف يحدث داخل الشبكة.

This comes at the cost of model interpretability, since as you add in more hidden layers the neurons start representing more and more abstract and ultimately unintelligible features — to the point that you may hear deep learning referred to as “black box optimization”, where you basically are just trying stuff somewhat at random and seeing what comes out, without really understanding what’s happening inside.



Machine Learning for Humans

. الحسر الخطي قابل للتفسير لأنك قررت الميزات التي يجب تضمينها في النموذج. يصعب تفسير الشبكات العصبية العميقة لأن الميزات يتم تعلمها ولا يتم شرحها في أي مكان باللغة العربية. كل شيء في مخيلة الآلة.

بعض الامتدادات والمفاهيم الإضافية التي تستحق الذكر

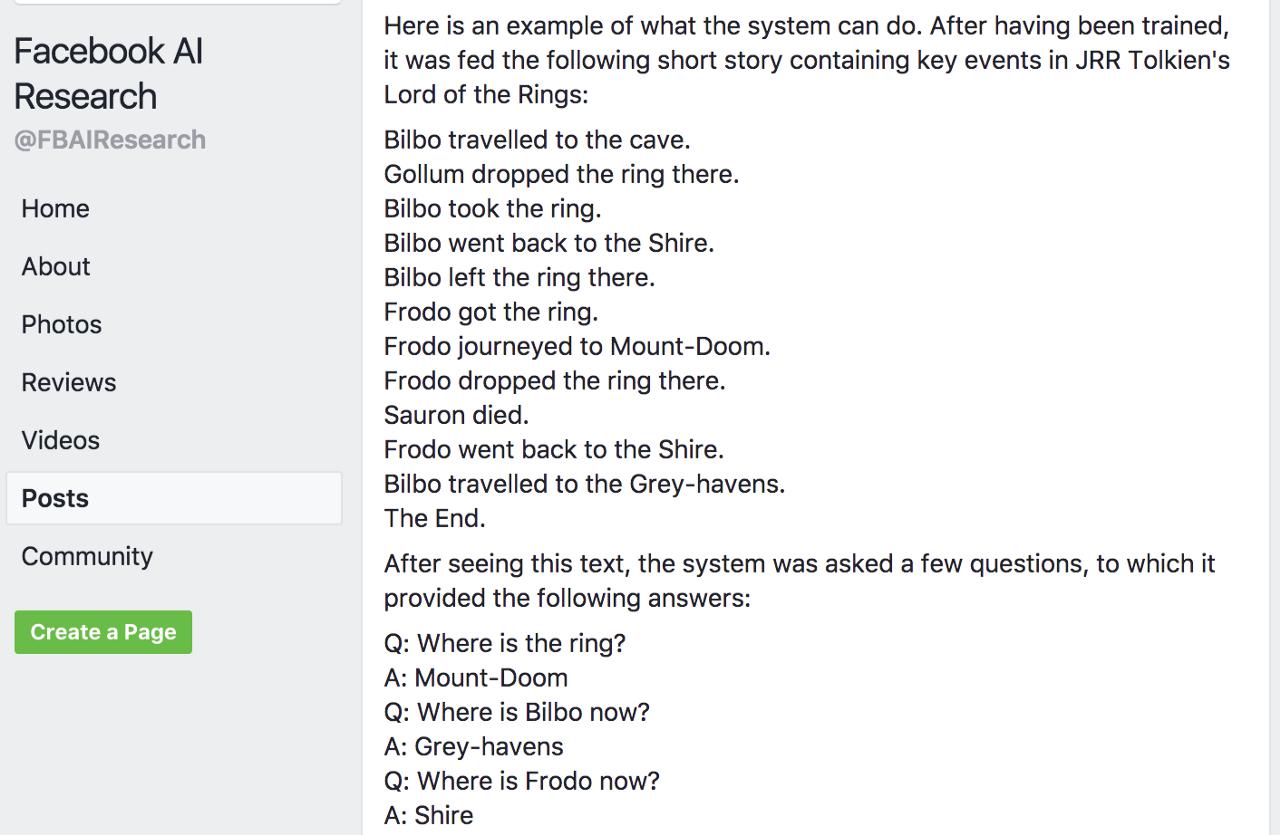
* **حزم برامج التعلم العميق**. نادراً ما تحتاج إلى تنفيذ جميع أجزاء الشبكات العصبية من البداية نظرًا للمكتبات والأدوات الموجودة التي تجعل تطبيقات التعلم العميق أسهل. هناك العديد من هذه: TensorFlow ، Caffe ، Torch ، Theano ، وأكثر من ذلك.
* • الشبكات العصبية التلافيفية . (شبت) تم تصميم (شبت) خصيصا لالتقاط الصور كمدخلات ، وفعالة لمهام رؤية الحاسوب. كما أنها مفيدة في التعلم المقوى العميق. مستوحاة على وجه التحديد من طريقة عمل القشرة البصرية الحيوانية ، وهي محور دورة التعلم العميق التي كنا نشير إليها في جميع أنحاء هذا المقال ، CS231n في جامعة ستانفورد.
* الشبكات العصبية المتكررة (شبر) . لدى (شبر) شعور بالذاكرة المدمجة وهي مناسبة تمامًا لمشاكل اللغة. كما أنها مهمة في التعلم المقوى نظرًا لأنها تمكن المنظومة من تتبع مكان وجود الأشياء وما حدث تاريخيًا حتى عندما لا تكون كل هذه العناصر مرئية في وقت واحد. [كتب كريستوفر](http://colah.github.io/posts/2015-08-Understanding-LSTMs/) أولاه مجموعة ممتازة من RNNs و LSTMs في سياق مشاكل اللغة.
* **Deep reinforcement learning.** This is one of the most exciting areas of deep learningresearch, at the heart of recent achievements like OpenAI defeating professional Dota 2 players and DeepMind’s AlphaGo surpassing humans in the game of Go. We’ll dive deeper in [Part 5](https://medium.com/@v_maini/machine-learning-for-humans-part-5-reinforcement-learning-6eacf258b265), but essentially the goal is to apply all of the techniques in this post to the problem of teaching an agent to maximize reward. This can be applied in any context that can be gamified — from actual games like Counter Strike or Pacman, to self-driving cars, to trading stocks, to (ultimately) real life and the real world.

Machine Learning for Humans

تطبيقات تعلم العميق

Deep learning is reshaping the world in virtually every domain. Here are a few examples of the incredible things that deep learning can do…

* *Facebook trained a neural network augmented by short-term memory to intelligently answer questions about the plot of Lord of the Rings.*



Research from [FAIR](https://www.facebook.com/FBAIResearch/posts/362517620591864) (Facebook AI Research) applying deep neural networks augmented by separate short-term memory to intelligently answer questions about the LOTR storyline. This is the definition of epic.

* *Self-driving cars rely on deep learning for visual tasks like understanding road signs, detecting lanes, and recognizing obstacles.*



Source: [Business Insider](http://www.businessinsider.com/uber-driverless-car-in-pittsburgh-review-photos-2016-9/#to-try-the-cars-we-lined-up-at-ubers-advanced-technologies-center-in-the-strip-district-of-pittsburgh-a-small-neighborhood-on-the-allegheny-river-with-nearby-warehouses-the-atc-is-tucked-under-an-overpass-for-a-freight-train-keeping-it-secluded-1)

* *Deep learning can be used for fun stuff like art generation. A tool called* [neural](https://github.com/jcjohnson/neural-style)[style](https://github.com/jcjohnson/neural-style) *can* *impressively mimic an artist’s style and use it to remix another image.*



The style of Van Gogh’s [Starry Night](https://en.wikipedia.org/wiki/The_Starry_Night) applied to a picture of Stanford’s campus, via Justin Johnson’s neural style

implementation: <https://github.com/jcjohnson/neural-style>

Other noteworthy examples include:

* *Predicting molecule bioactivity for* [drug discovery](https://arxiv.org/abs/1510.02855)
* *Face and object recognition for photo and video tagging*
* *Powering Google search results*
* *Natural language understanding and generation, e.g.* [Google Translate](https://translate.google.com/)
* *The Mars explorer robot Curiosity is* [autonomously selecting inspection-worthy](https://www.theatlantic.com/technology/archive/2017/06/mars-curiosity-rover/531339/)

[soil targets](https://www.theatlantic.com/technology/archive/2017/06/mars-curiosity-rover/531339/) *based* *on visual examination* …and many, many, more.

Now go do it!

We haven’t gone into as much detail here on how neural networks are set up in practice because it’s much easier to understand the details by implementing them yourself. Here are some amazing hands-on resources for getting started.

* Play around with the architecture of neural networks to see how different configurations affect network performance with the Google’s [Neural Network Playground](http://playground.tensorflow.org/#activation=tanh&batchSize=10&dataset=circle&regDataset=reg-plane&learningRate=0.03&regularizationRate=0&noise=0&networkShape=4,2&seed=0.70277&showTestData=false&discretize=false&percTrainData=50&x=true&y=true&xTimesY=false&xSquared=false&ySquared=false&co).
* Get up-and-running quickly with this tutorial by Google: [TensorFlow and deep learning,](https://codelabs.developers.google.com/codelabs/cloud-tensorflow-mnist/) [without a PhD](https://codelabs.developers.google.com/codelabs/cloud-tensorflow-mnist/). Classify handwritten digits at >99% accuracy, get familiar with TensorFlow, and learn deep learning concepts within 3 hours.
* بعد ذلك ، حاول ان تحل اول واجب من سلسلة محاضرات برنستون (CS231n) , هدف الواجب هو بناء شبكة عصبية من طبقتين . الهدف هو ترسيخ المفاهيم.

Further resources

التعلم العميق هو موضوع موسع. وفقًا لذلك ، قمنا أيضًا بتجميع بعض أفضل الموارد التي واجهناها حول هذا الموضوع ، في حال كنت ترغب في التعمق ... أكثر.

* [Deeplearning.ai](http://deeplearning.ai), Andrew Ng’s new deep learning course with a comprehensive syllabus on the subject
* [CS231n: Convolutional Neural Networks for Visual Recognition](http://cs231n.stanford.edu/syllabus.html), Stanford’s deep learning course. One of the best treatments we’ve seen, with excellent lectures and illustrative problem sets
* Deep Learning & Neural Networks — accessible but rigorous
* [Deep Learning Book](http://www.deeplearningbook.org/)—foundational, more mathematical
* [Fast.ai](http://fast.ai)— less theoretical, much more applied and black-boxy
* See Greg Brockman (CTO of OpenAI)’s answer to the question “What are the best ways to pick up Deep Learning skills as an engineer?” on [Quora](https://www.quora.com/What-are-the-best-ways-to-pick-up-Deep-Learning-skills-as-an-engineer/answer/Greg-Brockman?srid=2sq8)

Next up: time to play some games!

Last, but most certainly not least, is [Part 5: Reinforcement Learning](https://medium.com/@v_maini/reinforcement-learning-6eacf258b265).



80

Part 4: Neural Networks & Deep Learning

الجزء الخامس تعلم المقوى (تم)



*الأستكشاف والأستغلال. قرارات عملية ماركوف . تعلم-كيو . تعلم سياسة و تعلم العميق المقوى*

“تناول بعض الحلوى كمكافئة لأنهائك الفصل سابق.”

في تعليم المشرف , تأتي بينانات تدريب مع الأجابات من "المشرف" ويا ليت الحياة جائت بأجاباهتها

في (تم) لا يوجد اجابة ولكن يقوم المتعلم بتخاذ القرار بناء على كيفية اداء المهمة . في ظل غياب بيانات تدريب . يقوم المتعلم بتعلم من الخبرة . حيث يقوم بجمع امثلة تدريبية (هذا الفعل كان جيد , هذا الفعل كان سيء) عبر طريق تجربة و الخطاء . بينما يقوم بمحاولة تعظيم الجائزة طويلة الامد .

سنسكتشف في هذا الفصل:

* *مقايضة الاستكشاف / الاستغلال*
* *عمليات قرار ماركوف (MDPs) البيئة الكلاسيكي لمهام RL*
* ***تعلم-كيو ، تعلم السياسة و تعلم المقوى العميق***
* ***وأخيراً ، مشكلة تعلم القيمة***

لنضع فأرً اليآً في متاهة

إن أسهل سياق للتفكير حول (تم) هو في ألعاب ذات هدف واضح ونظام نقطة.

Machine Learning for Humans

لنفترض أننا نلعب لعبة يسعي فيها الفار إلى الحصول على أقصى مكافأة من الجبن في نهاية المتاهة (+1000 نقطة) ، أو مكافأة الماء الأقل على طول الطريق (+10 نقاط). في الوقت نفسه ، يريد الفار تجنب المواقع التي تقدم صدمة كهربائية (-100 نقطة).



المكافئة هي الجبن.

بعد قليل من الاستكشاف ، قد يجد الفأر واحة مصغرة وهي عبارة عن ثلاثة مصادر مائية متجمعة بالقرب من المدخل ، ويقضي كل وقته في استغلال هذا الاكتشاف من خلال الحصول على المكافآت صغيرة لمصادر المياه هذه باستمرار وعدم الذهاب إلى مزيد من المتاهة لمتابعة الجائزة الأكبر.

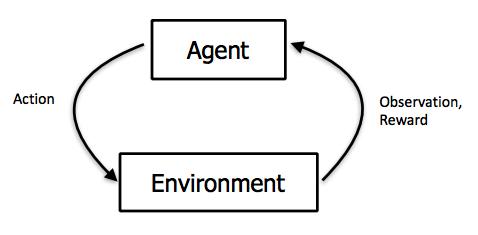
ولكن كما ترون ، فإن الفأر سيفوت على واحة أفضل في المتاهة ، أو المكافأة النهائية للجبن في النهاية!

هذا يحصر مقايضة استغلال / استكشاف. تتمثل إحدى الاستراتيجيات البسيطة للاستكشاف في أن يتخذ الفار أفضل إجراء معروف في معظم الأوقات (على سبيل المثال ، 80٪ من الوقت) ، ولكن في بعض الأحيان يستكشف اتجاهًا جديدًا تم اختياره عشوائيًا على الرغم من أنه قد يكون بعيدًا عن المكافأة المعروفة.

Machine Learning for Humans

وتسمى هذه الاستراتيجية استراتيجية (طمع ابسلون) ، حيث ابسلون هي النسبة المئوية من الوقت الذي يتخذ فيه العميل إجراءً منتقاة عشوائياً بدلاً من اتخاذ الإجراء الذي يُرجح أن يزيد المكافأة في ضوء ما يعرفه حتى الآن (في هذه الحالة ، 20 ٪). نبدأ عادة بالكثير من الاستكشاف (أي قيمة أعلى ل). وبمرور الوقت ، عندما يتعلم الفأر المزيد عن المتاهة وأيها يحقق أفضل مكافأة على المدى الطويل ، سيكون من المنطقي خفض ابسلون بثبات إلى 10٪ أو حتى أقل نظرًا لاستقراره في استغلال ما يعرفه.

من المهم أن تضع في اعتبارك أن المكافأة ليست فورية دائمًا: ففي مثال الروبوت - على سبيل المثال ، قد يكون هناك امتداد طويل من المتاهة يجب عليك المرور خلالها والعديد من نقاط القرار قبل الوصول إلى الجبن.



The agent observes the environment, takes an action to interact with the environment, and receives positive or negative reward. Diagram from Berkeley’s [CS 294: Deep Reinforcement Learning](CS%20294:%20Deep%20Reinforcement%20Learning) by John Schulman & Pieter Abbeel

قرارات عملية ماركوف (رما)

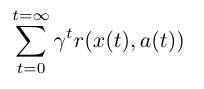
يمكن إضفاء الطابع الرسمي على تجوال الفأر عبر المتاهة كعملية قرار ماركوف ، وهي العملية التي تحدد احتمالات الأنتقال من حالة إلى حالة .سنشرح ذلك بالإشارة إلى مثالنا على الفأر الألي :

1. *مجموعة محدودة من الحالات. هذه هي المواقف المحتملة للماوس لدينا داخل المتاهة..*
2. ***مجموعة من الإجراءات المتاحة في كل ولاية. هذا هو {للأمام , للخلف} في ممر و {للأمام , للخلف , لليمين , لليسار} في تقاطع طرق***

Machine Learning for Humans

1. ***Transitions between states.*** *For example, if you go left at a crossroads**you end up in a new position. These can be a set of probabilities that link to more than one possible state (e.g. when you use an attack in a game of Pokémon you can either miss, inflict some damage, or inflict enough damage to knock out your opponent).*
2. ***Rewards associated with each transition.*** *In the robot-mouse example,**most of the rewards are 0, but they’re positive if you reach a point that has water or cheese and negative if you reach a point that has an electric shock.*
3. ***A discount factor γ between 0 and 1.*** *This quantifies the difference in**importance between immediate rewards and future rewards. For example, if* γ *is .9, and there’s a reward of 5 after 3 steps, the present value of that reward is .9³\*5.*
4. ***Memorylessness.*** *Once the current state is known, the history of the**mouse’s travels through the maze can be erased because the current Markov state contains all useful information from the history. In other words, “the future is independent of the past given the present”.*

Now that we know what an MDP is, we can formalize the mouse’s objective. We’re trying to maximize the sum of rewards in the long term:



Let’s look at this sum term by term. First of all, we’re summing across all time steps t. Let’s set γ at 1 for now and forget about it. r(x,a) is a reward function. For state x and action a (i.e., go left at a crossroads) it gives you the reward associated with taking that action a at state x. Going back to our equation, we’re trying to maximize the sum of future rewards by taking the best action in each state.

Machine Learning for Humans

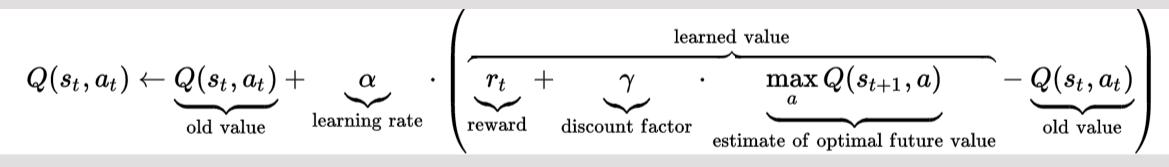
Now that we’ve set up our reinforcement learning problem and formalized the goal, let’s explore some possible solutions.

Q-learning: learning the action-value function

Q-learning is a technique that evaluates which action to take based on an **action-value** **function** that determines the value of being in a certain state and taking a certainaction at that state.

We have a function Q that takes as an input one state and one action and returns the **expected reward** of that action (and all subsequent actions) at that state. Before we explore the environment, Q gives the same (arbitrary) fixed value. But then, as we explore the environment more, Q gives us a better and better approximation of the value of an action a at a state s. We update our function Q as we go.

This equation from the [Wikipedia page on Q-learning](https://en.wikipedia.org/wiki/Q-learning) explains it all very nicely. It shows how we update the value of Q based on the reward we get from our environment:



Let’s ignore the discount factor γ by setting it to 1 again. First, keep in mind that Q is supposed to show you the full sum of rewards from choosing action Q and all the optimal actions afterward.

Now let’s go through the equation from left to right. When we take action at in state st, we update our value of Q(st,at) by adding a term to it. This term contains:

* *Learning rate* ***alpha****: this is how aggressive we want to be when updating our value. When alpha is close to 0, we’re not updating very aggressively. When alpha is close to 1, we’re simply replacing the old value with the updated value.*
* *The* ***reward is*** *the reward we got by taking action at at state st. So we’re adding*



85

Part 5: Reinforcement Learning

Machine Learning for Humans

*this reward to our old estimate.*

* *We’re also adding the* ***estimated future reward****, which is the maximum achievable reward Q for all available actions at xt+1.*
* *Finally, we subtract the old value of Q to make sure that we’re only incrementing or decrementing by the difference in the estimate (multiplied by alpha of course).*

Now that we have a value estimate for each state-action pair, we can select which action to take according to our **action-selection strategy** (we don’t necessarily just choose the action that leads to the most expected reward every time, e.g. with an epsilon-greedy exploration strategy we’d take a random action some percentage of the time).

In the robot mouse example, we can use Q-learning to figure out the value of each position in the maze and the value of the actions {forward, backward, left, right} at each position. Then we can use our action-selection strategy to choose what the mouse actually does at each time step.

Policy learning: a map from state to action

In the Q-learning approach, we learned a value function that estimated the value of each state-action pair.

**Policy learning** is a more straightforward alternative in which we learn a **policy function**,π, which is a direct map from each state to the best corresponding action at that state. Think of it as a behavioral policy: “when I observe state s, the best thing to

do is take action a”. For example, an autonomous vehicle’s policy might effectively include something like: “if I see a yellow light and I am more than 100 feet from the intersection, I should brake. Otherwise, keep moving forward.”



السياسة هي عبارة عن تحول من حالة إلى فعل



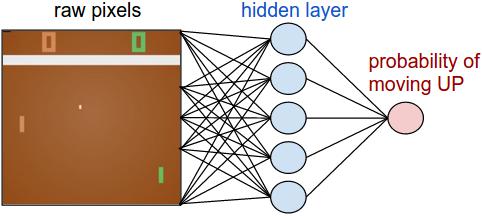
86

Part 5: Reinforcement Learning

Machine Learning for Humans

So we’re learning a function that will maximize expected reward. What do we know that’s really good at learning complex functions? Deep neural networks!

Andrej Karpathy’s [Pong from Pixels](http://karpathy.github.io/2016/05/31/rl/) provides an excellent walkthrough on using **deep** **reinforcement learning** to learn a policy for the Atari game Pong that takes raw pixelsfrom the game as the input (state) and outputs a probability of moving the paddle up or down (action).



In a policy gradient network, the agent learns the optimal policy by adjusting its weights through gradient descent based on reward signals from the environment. Image via <http://karpathy.github.io/2016/05/31/rl/>

If you want to get your hands dirty with deep RL, work through Andrej’s post. You will implement a 2-layer policy network in 130 lines of code, and will also learn how to plug into OpenAI’s [Gym](https://gym.openai.com/), which allows you to quickly get up and running with your first reinforcement learning algorithm, test it on a variety of games, and see how its performance compares to other submissions.

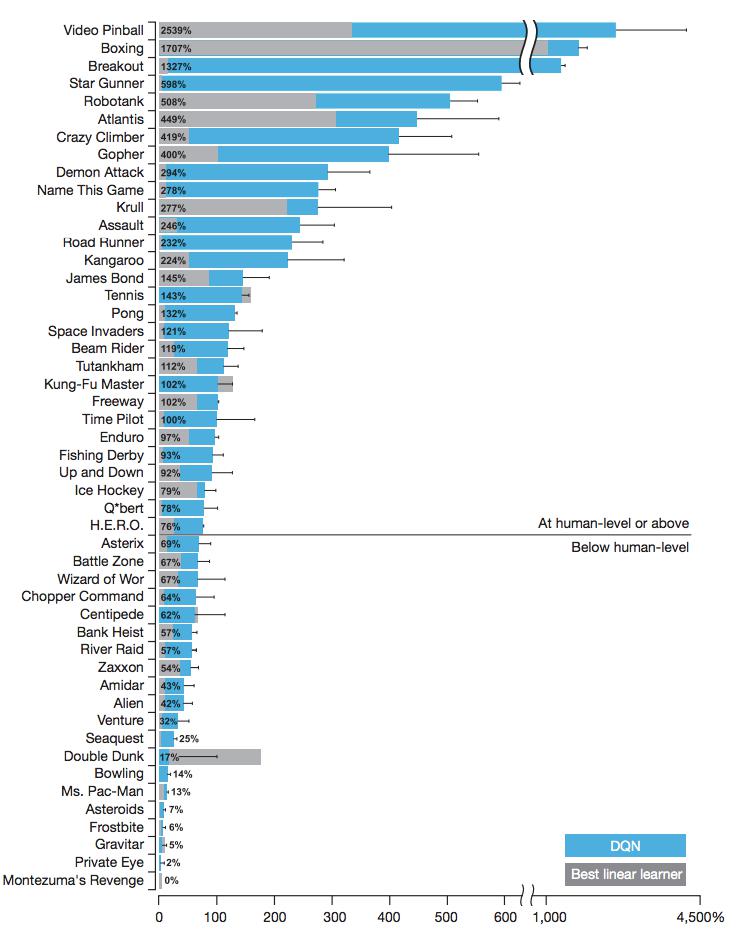
DQNs, A3C, and advancements in deep RL

In 2015, DeepMind used a method called deep Q-networks (DQN), an approach that approximates Q-functions using deep neural networks, to beat human benchmarks across many Atari games:

Machine Learning for Humans

*We demonstrate that the deep Q-network agent, receiving only the pixels and the game score as inputs, was able to surpass the performance of all previous algorithms and achieve a level comparable to that of a professional human games tester across a set of 49 games, using the same algorithm, network architecture and hyperparameters. This work bridges the divide between high-dimensional sensory inputs and actions, resulting in the first artificial agent that is capable of learning to excel at a diverse array of challenging tasks. (*[Silver et al., 2015](https://storage.googleapis.com/deepmind-media/dqn/DQNNaturePaper.pdf)*)*

Here is a snapshot of where DQN agents stand relative to linear learners and humans in various domains:



These are normalized with respect to professional human games testers: 0% = random play, 100% = human performance.

Source: DeepMind’s DQN paper, [Human-level control through deep reinforcement learning](https://storage.googleapis.com/deepmind-media/dqn/DQNNaturePaper.pdf)

Machine Learning for Humans

To help you build some intuition for how advancements are made in RL research, here are some examples of improvements on attempts at non-linear Q-function approximators that can improve performance and stability:

* **Experience replay***, which learns by randomizing over a longer sequence of**previous observations and corresponding reward to avoid overfitting to recent experiences. This idea is* [inspired by biological brains](https://deepmind.com/blog/ai-and-neuroscience-virtuous-circle/)*: rats traversing mazes, for example, “replay” patterns of neural activity during sleep in order to optimize future behavior in the maze.*
* **Recurrent neural networks (RNNs)** *augmenting DQNs. When an agent can**only see its immediate surroundings (e.g. robot-mouse only seeing a certain segment of the maze vs. a birds-eye view of the whole maze), the agent needs to remember the bigger picture so it remembers where things are. This is similar to how humans babies develop* [object permanence](https://en.wikipedia.org/wiki/Object_permanence)*to know things exist even if they leave the baby’s visual field. RNNs are “recurrent”, i.e.* [they allow information to](http://colah.github.io/posts/2015-08-Understanding-LSTMs/)[persist](http://colah.github.io/posts/2015-08-Understanding-LSTMs/) *on* *a longer-term basis. Here’s an impressive video of a deep recurrent* *Q-network (DQRN) playing Doom.*

**

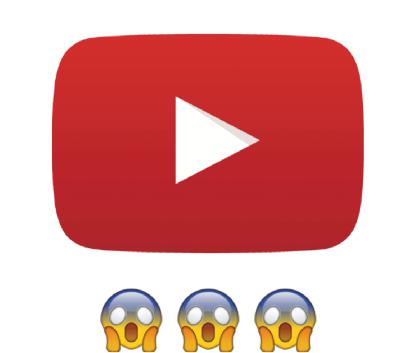
Paper: <https://arxiv.org/abs/1609.05521>. Source: Arthur Juliani’s [Simple Reinforcement Learning](https://medium.com/emergent-future/simple-reinforcement-learning-with-tensorflow-part-6-partial-observability-and-deep-recurrent-q-68463e9aeefc#.gi4xdq8pk) [with Tensorflow](https://medium.com/emergent-future/simple-reinforcement-learning-with-tensorflow-part-6-partial-observability-and-deep-recurrent-q-68463e9aeefc#.gi4xdq8pk) series

. في عام 2016 ، بعد عام واحد فقط من ورقة DQN ، كشفت DeepMind عن خوارزمية أخرى تسمى Asynchronous Advantage Actor-Critic (A3C) والتي تجاوزت الأداء الحديث في ألعاب Atari بعد التدريب لنصف المدة (Mnih et al. ، 2016 ). A3C عبارة عن خوارزمية ممثل - ناقد تجمع بين أفضل الطرق التي اكتشفناها سابقًا: فهي تستخدم جهة فاعلة (شبكة سياسة تقرر كيفية التصرف) وناقدًا (شبكة Q التي تقرر مدى قيمة الأشياء). كتب آرثر جولياني لطيفة عن كيفية عمل A3C على وجه التحديد. A3C هي الآن وكيل بداية الكون لـ OpenAI.

Machine Learning for Humans

منذ ذلك الحين ، كان هناك عدد لا يحصى من الاختراقات الرائعة - من اختراع الذكاء الاصطناعي لغتهم الخاصة إلى تعليم أنفسهم السير في مجموعة متنوعة من التضاريس. هذه السلسلة تخدش السطح فقط على حافة القطع لـ RL ، ولكن نأمل أن تكون بمثابة نقطة انطلاق لمزيد من الاستكشاف!

كملاحظة فاصلة ، نود مشاركة هذا الفيديو المذهل لعملاء DeepMind الذين تعلموا المشي ... بصوت إضافي. احصل على بعض الفشار وارفع مستوى الصوت وشاهد المجد الكامل للذكاء الاصطناعي.



Practice materials & further reading

**Code**

* *Andrej Karpathy’s* [Pong from Pixels](http://karpathy.github.io/2016/05/31/rl/)*will get you up-and-running quickly with your first reinforcement learning agent. As the article describes, “we’ll learn to play an ATARI game (Pong!) with PG, from scratch, from pixels, with a deep neural network, and the whole thing is 130 lines of Python only using numpy as a dependency (*[Gist link](https://gist.github.com/karpathy/a4166c7fe253700972fcbc77e4ea32c5)*).”*
* *Next, we’d highly recommend Arthur Juliani’s* [Simple Reinforcement Learning](https://medium.com/emergent-future/simple-reinforcement-learning-with-tensorflow-part-0-q-learning-with-tables-and-neural-networks-d195264329d0)[with Tensorflow](https://medium.com/emergent-future/simple-reinforcement-learning-with-tensorflow-part-0-q-learning-with-tables-and-neural-networks-d195264329d0) *tutorial.* *It walks through DQNs, policy learning, actor-critic* *methods, and strategies for exploration with implementations using TensorFlow. Try understanding and then re-implementing the methods covered.*

Reading + lectures

* *Richard Sutton’s book,* [Reinforcement Learning: An Introduction](http://people.inf.elte.hu/lorincz/Files/RL_2006/SuttonBook.pdf)*—a fantastic book, very readable*
* *John Schulman’s* [CS 294: Deep Reinforcement Learning](http://rll.berkeley.edu/deeprlcourse/)*(Berkeley)*
* *David Silver’s* [Reinforcement Learning](http://www0.cs.ucl.ac.uk/staff/d.silver/web/Teaching.html)*course (UCL*)

لقد فعلت ذلك!

*.* *إذا كنت قد وصلت إلى هذا الحد ، فهذا هو كل المكافأة التي نأمل أن نحصل عليها.*

*نأمل أن تكون قد استمتعت بهذه السلسلة كمقدمة للتعلم الآلي. لقد قمنا بتجميع بعض موارد تعلم الآلة المفضلة لدينا في الملحق إذا كنت مستعدًا لمعرفة مدى عمق حفرة الأرانب هذه.*

*!* *لا تتردد في التواصل مع الأفكار أو الأسئلة أو التعليقات أو صور GIF المفضلة لديك!*

*للمرة القادمة*

*فيشال وسامر*

افكار ختامية:

هناك سؤال أساسي استلهمنا منه هذه السلسلة ، ونود أن نطرحها لك أيضًا.

ما هي وظيفتنا الهدفة ، كبشر؟ كيف نحدد المكافأة التي نزيدها في حياتنا الحقيقية؟ بالإضافة إلى المتعة والألم الأساسيين ، فإن تعريفنا للمكافأة يميل أيضًا إلى تضمين أشياء فوضوية مثل الصواب والخطأ والأستيفاء والحب والروحانية والهدف.

لقد كان هناك مجال فكري مخصص لمسألة ما هي وظيفتنا الموضوعية أو ينبغي أن تكون منذ العصور القديمة ، ويسمى الفلسفة الأخلاقية. المسألة المركزية للفلسفة الأخلاقية هي: ما الذي يجب علينا فعله؟ كيف نعيش؟ ما هي الأعمال الصحيحة أم الخاطئة؟ الجواب ، بكل وضوح: إنه يعتمد على قيمك.

. مع قيامنا بإنشاء المزيد والمزيد من الذكاء الاصطناعي المتقدم ، سيبدأ في الابتعاد عن عالم مشاكل الألعاب مثل ألعاب آتاري ، حيث يتم تحديد "المكافأة" بوضوح من خلال عدد النقاط التي تم ربحها في اللعبة ، وتوجد أكثر فأكثر في العالم الحقيقي . على سبيل المثال ، يتعين على المركبات المستقلة اتخاذ قرارات بتعريف أكثر تعقيدًا إلى حد ما للمكافأة. في البداية ، قد يتم ربط المكافأة بشيء مثل "الوصول بأمان إلى الوجهة". ولكن إذا اضطررت للاختيار بين البقاء في المسار والاصطدام بخمسة مشاة أو الانحراف وضرب واحد ، فهل يجب أن تنحرف السيارة؟ ماذا لو كان أحد المشاة طفلًا ، أو مسلح طليق ، أو أينشتاين التالي؟ كيف يغير ذلك القرار ، ولماذا؟ ماذا لو أدى الانحراف أيضًا إلى تدمير قطعة فنية قيّمة؟ ، والإجابات ليست بهذه البساطة.

في هذه السلسلة ، اكتشفنا سبب صعوبة تحديد شكل قطة بشكل صريح لجهاز كمبيوتر - إذا سئلنا كيف نعرف أنفسنا ، فإن الإجابة هي ، ببساطة ، "الحدس" - لكننا اكتشفنا مناهج رؤية الآلة لتعليم آلة لتعلم هذا الحدس من تلقاء نفسها. وبالمثل ، في مجال أخلاق الآلة ، قد يكون من الصعب تحديد كيفية تقييم صواب أو خطأ فعل ما مقابل الآخر ، ولكن ربما يكون من الممكن للآلة أن تتعلم هذه القيم بطريقة ما. وهذا ما يسمى بمشكلة تعلم القيم ، وقد تكون واحدة من أهم المشكلات الفنية التي سيتعين على البشر حلها على الإطلاق.

لمزيد من المعلومات حول هذا الموضوع ، راجع هذا المنشور الشامل حول مخاطر الذكاء الاصطناعي. ومع تقدمك في عالم جعل الآلة أكثر ذكاءً وذكاءً ، فإننا نشجعك على أن تضع في اعتبارك أن تقدم الذكاء الاصطناعي هو سيف ذو حدين ، وحرص خاص على كلا الجانبين.

92

الملحق : افضل مصادر لتعلم حول تعلم الآلة



*.* *مجموعة من الموارد لصياغة منهج حول الذكاء الاصطناعي والتعلم الآلي والتعلم العميق.*

نصيحة عامة حول صياغة المنهج

إن الذهاب إلى للكلية للحصول على شهادة جامعية أمر غير ممكن أو مرغوب فيه دائمًا. بالنسبة لأولئك الذين يفكرون في بديل تلقائي ، هذا هو لك.

**1 – تعلم الآساسيات ثمة تخصص في المجال الذي ترغب به**

لا يمكنك تعمق في كافة مجلات تعلم الآلة . يوجد الكثير لتعلمه و المجال ينمو ويتطور بسرعة . اتقن الاساسيات ثمة ركز على مشاريع في المجال الذي يداعب اهتمامك – اذ كان ذالك معالجة اللغة طبيعية , رؤية الحاسوب , تعلم العميق المقوى , الآليات او اي شيء اخر.

**2 – صمم المنهج حول المواضيع التي تهمك**

الدافع هو أكثر أهمية بكثير من التحسين الجزئي لاستراتيجية التعلم لبعض الأهداف الأكاديمية أو المهنية طويلة الأجل. إذا كنت تستمتع ، فسوف تحقق تقدمًا سريعًا. إذا كنت تحاول إجبار نفسك على التقدم ، فستتباطأ.

Machine Learning for Humans

الآساسيات

البرمجة

Syntax and basic concepts: [Google’s Python Class](https://developers.google.com/edu/python/), [Learn Python the Hard Way](https://learnpythonthehardway.org/book/ex0.html).

Practice: [Coderbyte](https://coderbyte.com/), [Codewars](https://www.codewars.com/), [HackerRank](https://www.hackerrank.com/).

الجبر الخطي

[Deep Learning Book, Chapter 2: Linear Algebra](http://www.deeplearningbook.org/contents/linear_algebra.html). A quick review of the linear algebra concepts relevant to machine learning.

[A First Course in Linear Model Theory](https://www.amazon.com/First-Course-Linear-Model-Theory/dp/1584882476) by Nalini Ravishanker and Dipak Dey. Textbook introducing linear algebra in a statistical context.

الأحصاء و الأحتمالات

MIT 18.05, [Introduction to Probability and Statistics](https://ocw.mit.edu/courses/mathematics/18-05-introduction-to-probability-and-statistics-spring-2014/index.htm), taught by Jeremy Orloff and Jonathan Bloom. Provides intuition for probabilistic reasoning & statistical inference, which is invaluable for understanding how machines think, plan, and make decisions. [All of Statistics: A Concise Course in Statistical Inference](http://www.ic.unicamp.br/~wainer/cursos/1s2013/ml/livro.pdf), by Larry Wasserman. Introductory text on statistics.

تفضال و تكامل

Khan Academy: [Differential Calculus](https://www.khanacademy.org/math/calculus-home/differential-calculus). Or, any introductory calculus course or textbook.

Stanford CS231n: [Derivatives, Backpropagation, and Vectorization](http://cs231n.stanford.edu/handouts/derivatives.pdf), prepared by Justin Johnson.

تعلم الآلة

**دورات**

Andrew Ng’s [Machine Learning](https://www.coursera.org/learn/machine-learning) course on Coursera (or, for more rigor, [Stanford CS229](http://cs229.stanford.edu/materials.html)). Machine learning bootcamps: [Galvanize](https://www.galvanize.com/san-francisco/data-science) (full-time, 3 months, $$$$), [Thinkful](https://www.thinkful.com/bootcamp/data-science/flexible/) (flexible schedule, 6 months, $$).

**كتب**

[An Introduction to Statistical Learning](http://www-bcf.usc.edu/~gareth/ISL/) by Gareth James et al. Excellent reference for essential machine learning concepts, available free online.

Machine Learning for Humans

تعلم العميق

**الدورات**

[Deeplearning.ai](http://deeplearning.ai), Andrew Ng’s introductory deep learning course.

[CS231n: Convolutional Neural Networks for Visual Recognition](http://cs231n.stanford.edu/syllabus.html), Stanford’s deep learning course. Helpful for building foundations, with engaging lectures and illustrative problem sets.

**المشاريع**

[Fast.ai](http://fast.ai), a fun and hands-on project-based course. Projects include classifying images of dogs vs. cats and generating Nietzschean writing.

[MNIST handwritten digit classification with TensorFlow](https://codelabs.developers.google.com/codelabs/cloud-tensorflow-mnist/#0). Classify handwritten digits with >99% accuracy in 3 hours with this tutorial by Google.

Try your hand at [a Kaggle competition](https://www.kaggle.com/). Implement a deep learning paper that you found interesting, using other versions on GitHub as reference material.

**المطبوعات**

[Deep Learning Book](http://www.deeplearningbook.org/), a.k.a. the Bible of Deep Learning, authored by Ian Goodfellow, Yoshua Bengio, and Aaron Courville.

[Neural Networks and Deep Learning](http://neuralnetworksanddeeplearning.com/chap1.html), a clear and accessible online deep learning text by Michael Nielsen. Ends with commentary on reaching human-level intelligence. [Deep Learning Papers Reading Roadmap](https://github.com/songrotek/Deep-Learning-Papers-Reading-Roadmap), a compilation of key papers organized by chronology and research area.

تعلم المقوى

**الدورات**

John Schulman’s [CS 294: Deep Reinforcement Learning](http://rll.berkeley.edu/deeprlcourse/) at Berkeley.

David Silver’s [Reinforcement Learning](http://www0.cs.ucl.ac.uk/staff/d.silver/web/Teaching.html) course at University College London.

[Deep RL Bootcamp](https://www.deepbootcamp.io/), organized by OpenAI and UC Berkeley. Applications are currently closed, but it’s worth keeping an eye out for future sessions.

**مشاريع**

Andrej Karpathy’s [Pong from Pixels](http://karpathy.github.io/2016/05/31/rl/). Implement a Pong-playing agent from scratch in 130 lines of code.

Machine Learning for Humans

Arthur Juliani’s [Simple Reinforcement Learning with Tensorflow](https://medium.com/emergent-future/simple-reinforcement-learning-with-tensorflow-part-0-q-learning-with-tables-and-neural-networks-d195264329d0) series. Implement Q-learning, policy-learning, actor-critic methods, and strategies for exploration using TensorFlow.

See OpenAI’s [requests for research](https://openai.com/requests-for-research/) for more project ideas.

**المطبوات**

Richard Sutton’s book, [Reinforcement Learning: An Introduction](http://people.inf.elte.hu/lorincz/Files/RL_2006/SuttonBook.pdf).

الذكاء الاصطناعي

[Artificial Intelligence: A Modern Approach](http://aima.cs.berkeley.edu/) by Stuart Russell and Peter Norvig.

Sebastian Thrun’s Udacity course, [Intro to Artificial Intelligence](https://www.udacity.com/course/intro-to-artificial-intelligence--cs271).

Fellowships: [Insight AI Fellows Program](http://insightdata.ai/), [Google Brain Residency Program](https://research.google.com/teams/brain/residency/)

سلامة الذكاء الاصطناعي

For the short version, read: (1) Johannes Heidecke’s [Risks of Artificial Intelligence](https://thinkingwires.com/posts/2017-07-05-risks.html), (2)

OpenAI and Google Brain’s collaboration on [Concrete Problems in AI Safety](https://blog.openai.com/concrete-ai-safety-problems/), and (3)

Wait But Why’s article on the [AI Revolution](https://waitbutwhy.com/2015/01/artificial-intelligence-revolution-1.html).

For the longer version, see Nick Bostrom’s [Superintelligence](https://www.amazon.com/Superintelligence-Dangers-Strategies-Nick-Bostrom/dp/0198739834).

Check out the research published by the [Machine Intelligence Research Institute](https://intelligence.org/research/) (MIRI) and [Future of Humanity Institute](https://www.fhi.ox.ac.uk/research/research-areas/) (FHI) on AI safety. Keep up-to-date with [/r/ControlProblem](https://www.reddit.com/r/ControlProblem/) on Reddit.

مقالات

[Import AI](https://jack-clark.net/import-ai/), weekly AI newsletter covering the latest developments in the industry.

Prepared by Jack Clark of OpenAI.

[Machine Learnings](https://machinelearnings.co/), prepared by Sam DeBrule. Frequent guest appearances from experts in the field.

[Nathan.ai](http://nathan.ai), covering recent news and commenting on AI/ML from a venture capital perspective.

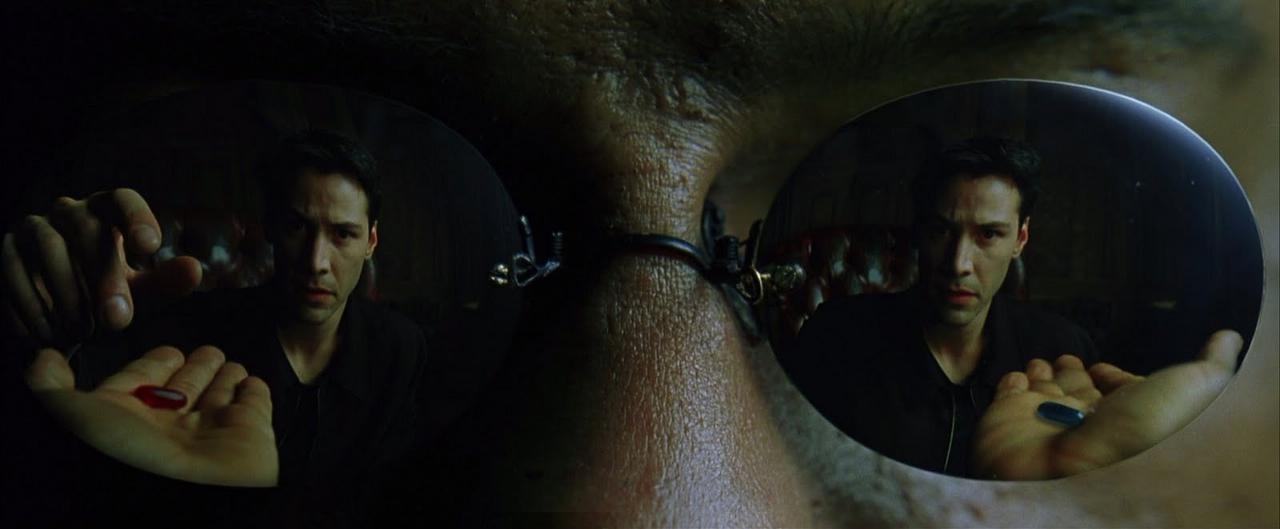
Machine Learning for Humans

نصائح

[“What is the best way to learn machine learning without taking any online courses?](https://www.forbes.com/sites/quora/2017/03/22/what-is-the-best-way-to-learn-machine-learning-without-taking-any-online-courses/#30fc6e5d5d87) — answered by Eric Jang, Google Brain

[What are the best ways to pick up deep learning skills as an engineer?"](https://www.quora.com/What-are-the-best-ways-to-pick-up-Deep-Learning-skills-as-an-engineer) - answered by Greg Brockman, CTO of OpenAI

A16z's [AI Playbook](http://aiplaybook.a16z.com/), a more code-based introduction to AI [AI safety syllabus](https://80000hours.org/ai-safety-syllabus/), designed by 80,000 Hours



"تأخذ الحبة الزرقاء ، تنتهي القصة. تستيقظ في سريرك وتؤمن بكل ما تريد أن تصدقه. تأخذ الحبة الحمراء ، وتبقى في بلاد العجائب ، وأريكم مدى عمق حفرة الأرانب ". - مورفيوس