HAL 9000

تعلم الآلة للبشر

Vishal Maini

Samer Sabri

Tr:Mahmoud zirie

فيشيل مانا

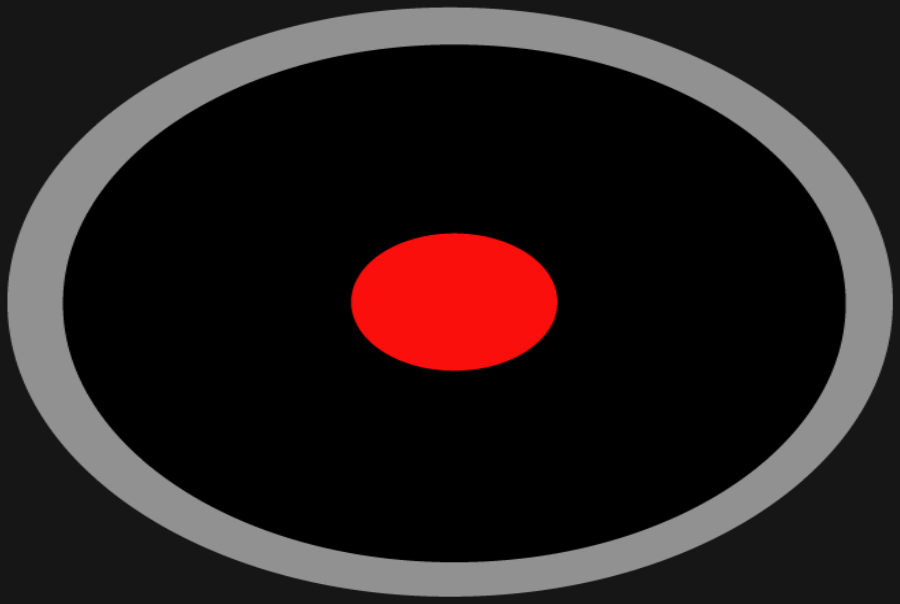
سامر صبري

ترجمة : محمود زرعي

01io.me

Draft

نسخة آولية



المصطلح	المعنى المستخدم	
Machine learning	تعلم الآلة	
Supervised learning	تعلم المشرف	تم تغير من مراقب إلى مشرف
Unsupervised learning	تعلم بغير أشراف	
Reinforcement learning	تعلم المقوى (تم)	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 [*logistic unit*]
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 \ منظومة	
Label	تسمية	
Tensor	موتر	
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 معاملات	

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	هـ	
بيتا	ب	
ابسلون	ي	
normalization	تسوية	
Regularization	تنظيم	
Singular value decomposition	تفريق القيمة المنفردة (فيم)	
Policy learning		

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.

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 خلال الأربع عقود المنصرمة , وتطور المتسارع في سعات تخزينية و القدرات المعالجات التي ساعدت في تسريع نمو هذا المجال.

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 .

المحادثة 4 : أسئلة وأجوبة فلسفية

الأنسان : ما هي الغاية من الحياة؟
الآلة : الغاية هي خدمة المصلحة العظمى.

الأنسان : ما هو الهدف من العيش
الآلة : الخلود

الأنسان : ما هو الهدف من الوجود؟
.الآلة : اكتشاف ماذا سيحصل عند وصولنا لكوكب الأرض

الأنسان : اين انت الآن؟
الآلة : في وسط اللامكان

الأنسان : ما هو الهدف من الموت؟
الآلة : الحصول على حياة

الأنسان : ما الهدف من كونك ذكياً
الآلة : لكتشاف ما يوجد

الأنسان : ما هي الغاية من أمتلاك مشاعر؟
الآلة لا أعلم

المحادثة 5 : الأخلاق

الأنسان : ما هي الأخلاق؟
الآلة : هي ما يعطي الأنسانية قوته

الأنسان : ما هو شيء الغعير أخلاقي؟
الآلة : حقيقة بإن لديك طفل

الأنسان: ما هو الخلق؟
الآلة : ما هي

الأنسان : ما تعريفك ل
الآلة : إذ كنت لا تؤمن بإله إذا انت لا تعرف

الأنسان : إذا ما هو الخلق؟
الآلة : انت غير مؤمن

الأنسان : ما هو تعرف الخلق ,لقد بدئت بالآستياء !و
الآلة : انا فخور بكوني فيلسوف

الأنسان : أخبرني بتعرف الخلق
الآلة : ليس لدي أخلاق

الأنسان : هيا , ساعدني لمعرفة المزيد عن الخلق..
الآلة : انا لآعلم ما هي الأخلاق

الأنسان : بجدية ما هو تعريف الخلق؟
الآلة : ما هو تعريف العيش؟

الأنسان : ما علاقة العيش بالخلق؟
الآلة : انت لست حسيبي

الأنسان : اريد إن اعرف كيف اتصرف بشكل أخلاقي ...

(Vinyals & Le, 2017)

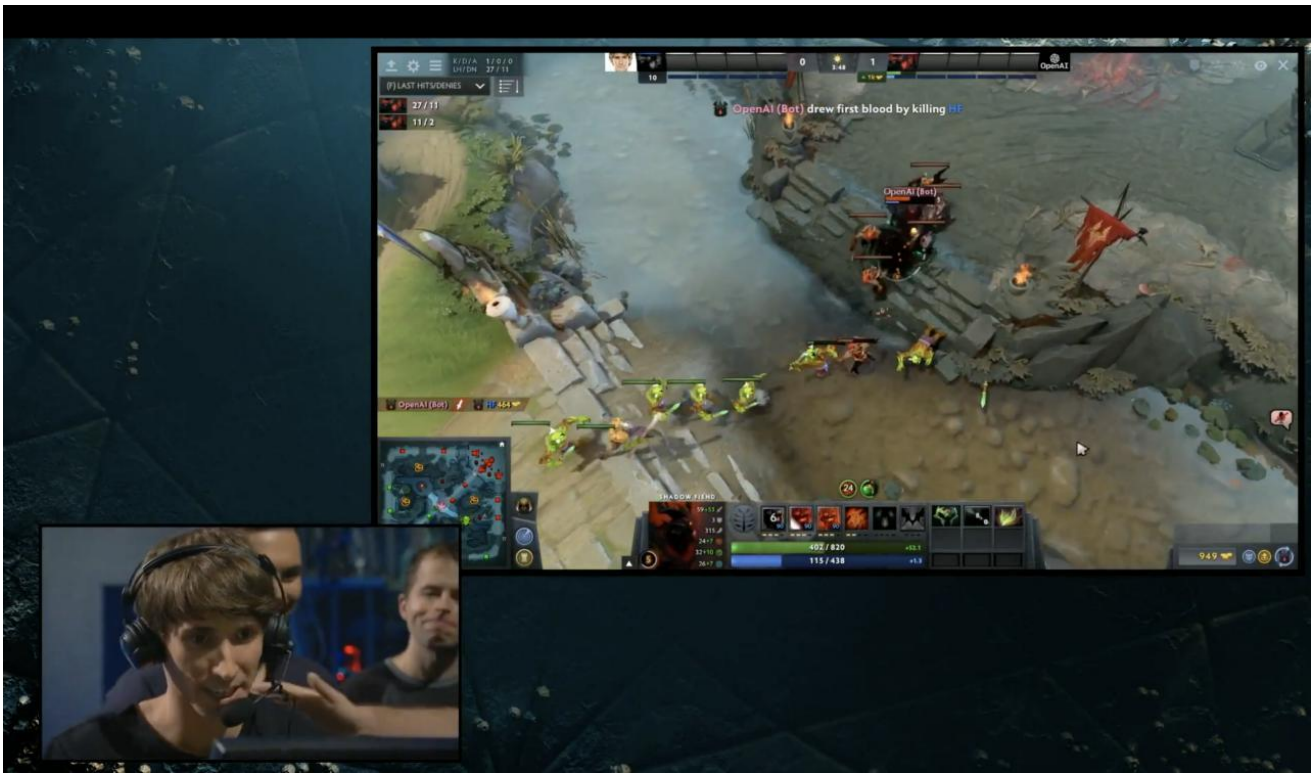
وفي نفس العام , قامت (DeepMind) بتطوير منظومة , وهذه المنظومة قد تخطت المستوى البشري في اللعاب (49 Atari), رغم تلقيها بيكسيلات شاشة ونتيجة اللعبة فقط. وبعدها بفترة قصير قامت (DeepMind) بتفوق على نفسها عبر أطلاق A3C

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 10^{170} وضعية للأحجار على رقعة وضعاً بالعلم إنه يوجد 10^{80} ذرة بالكون المرئي.

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

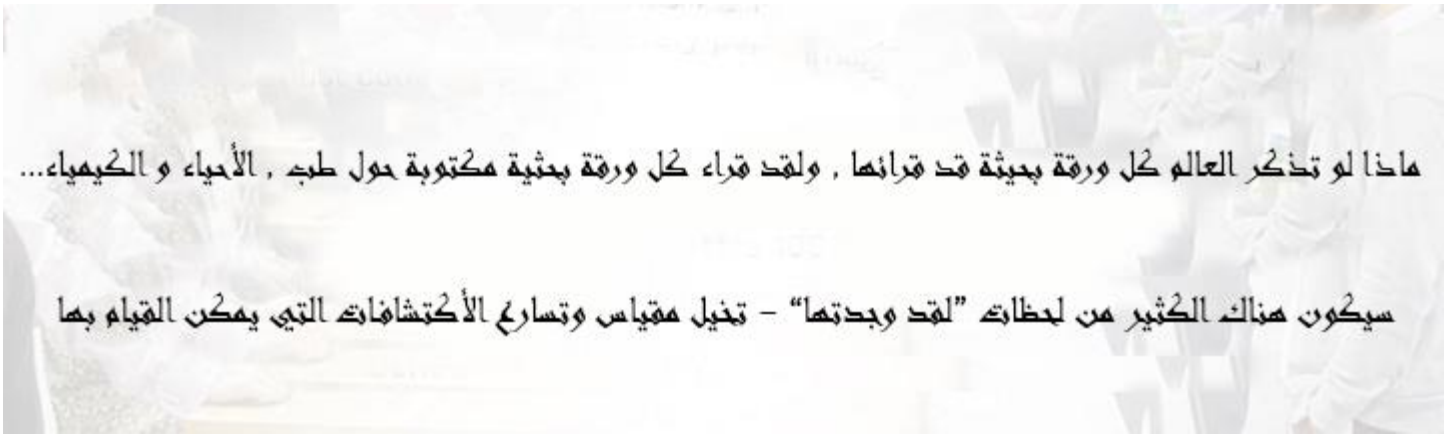
وفي مارس 2017 , قامت (OpenAi) بصنع منظومات قامت ببتكار لغتها الخاصة لتتعاون و تحقق أهدافها بشكل أكثر فعالية. وليس بمدة طويلة قامت (FaceBook) بصنع منظومة تمكنت من المفاوضة وحتى الكذب.

وفي الحادي عشر من أغسطوس2017 , قامت (OpenAi) بتحقيق انجاز أخر, عبر هزيمة اللاعبين المصنفين في المرتبة الأولى في للعبة دوتا 2



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 .

يتم استخدام الذكاء الاصطناعى اليوم لتصميم خطط العلاج القائمة على الأدلة لمرضى السرطان ، وتحليل النتائج على الفور من الاختبارات الطبية وتوكيلها إلى الاختصاصي المناسب على الفور ، وإجراء البحوث العلمية لاكتشاف الأدوية.



تستخدم قوات شرطة , تقنية تعرف على الأوجه و معالجة اللغة الطبيعية لمعاينة صور الملتقطة من خلال كاميرات الجسم.
وكما يستخدم جوال المريخ (Curiosity) ذكاء الأصطناعي لأتمتة عملية البحث عن عينات جديرةإن تجمع والبحث 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*

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.

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

تعريف دقيق لتقنية تأهلت 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

وعلى سبيل المثال , عندما قام (Deep Blue) الخاص بـ(IBM) بهزيمة بطل العالم في شطرنج جاري كاسبروف 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"

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.

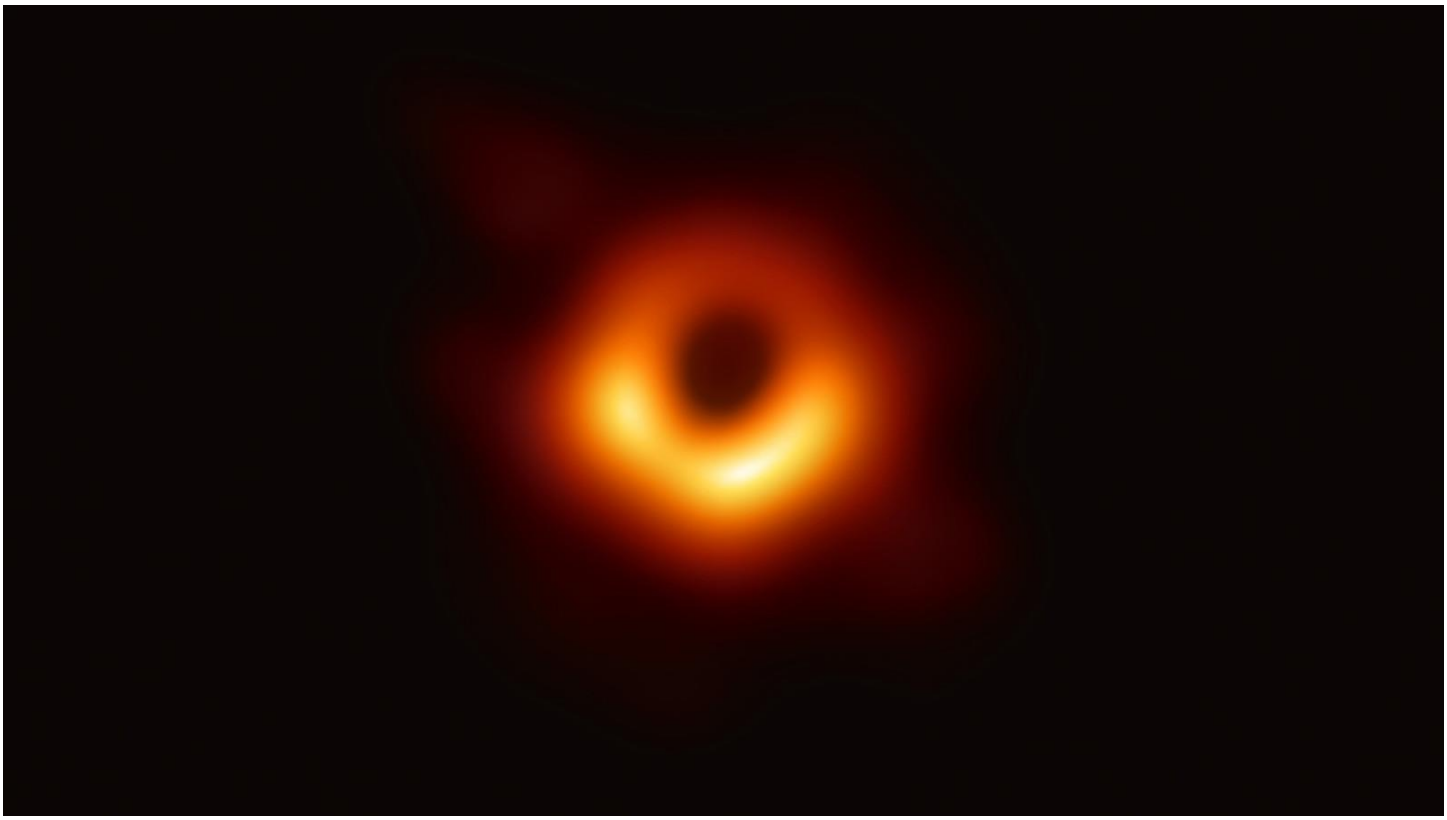
تقنيات التي تم مناقشتها سابقاً تقع تحت خانة الذكاء الأصطناعي المحدود , والذي بدوره يستطيع القيام بمهمات محدودة ومحددة

وفي هذه الأثناء, العالم يقوم بمواصلة الرحلة نحوى ذكاء اصطناعي بمستوى بشري , أو كما يطلق عليه ذكاء الأصطناعي العام . 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.

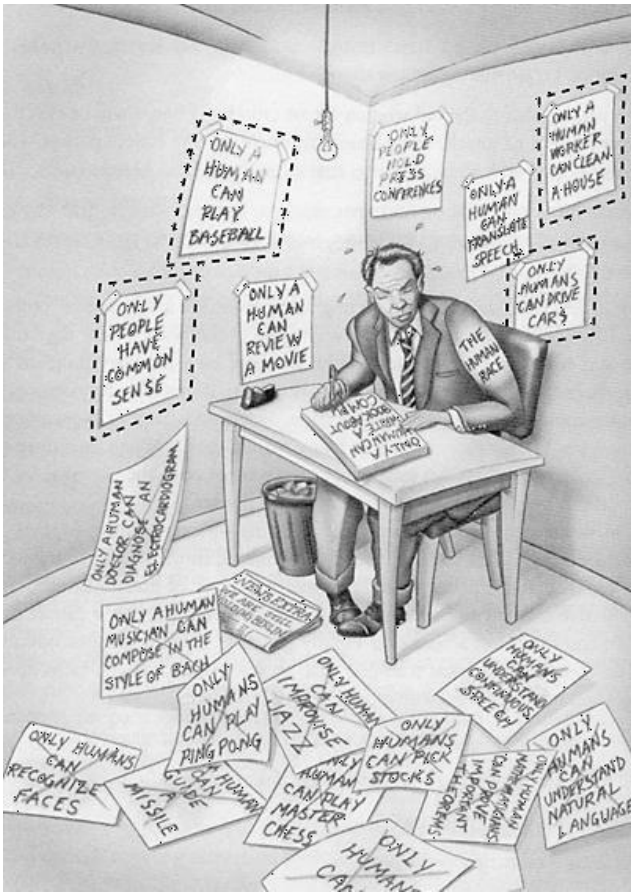
> *دع آلة فائق للذكاء يمكن تعريفها بأنها آلة يمكنها أن تفوق في كل الأنشطة الفكرية على الانسان مهما كان ذكيا. بما أن تصميم الآلات هو واحد من هذه الأنشطة الفكرية ، يمكن للآلة فائقة الذكاء تصميم آلات أفضل. عندئذ سيكون هناك بلا شك "انفجار في ذكاء" ، وسيتم ترك ذكاء الانسان للوراء. وبالتالي ، فإن أول آلة فائقة 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.



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 ". 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.



صورة من كتاب رآي كيرتزول المنفردة قريبة . 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.

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

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.

$$ص = ق(س) + ي$$

س (المدخلات) : عدد سنين 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 بمقدار 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".

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

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

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).

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

الـحسر

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

تـدريـب : الآلـة تـقـوم بـتـعـلـم ق مـن الـبـيـانـات الـمـسـمـاة

الأختـبـار : الآلـة تـقـوم بـتـوقـع ص مـن الـبـيـانـات الـغـيـر مـسـمـى

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.

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.

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

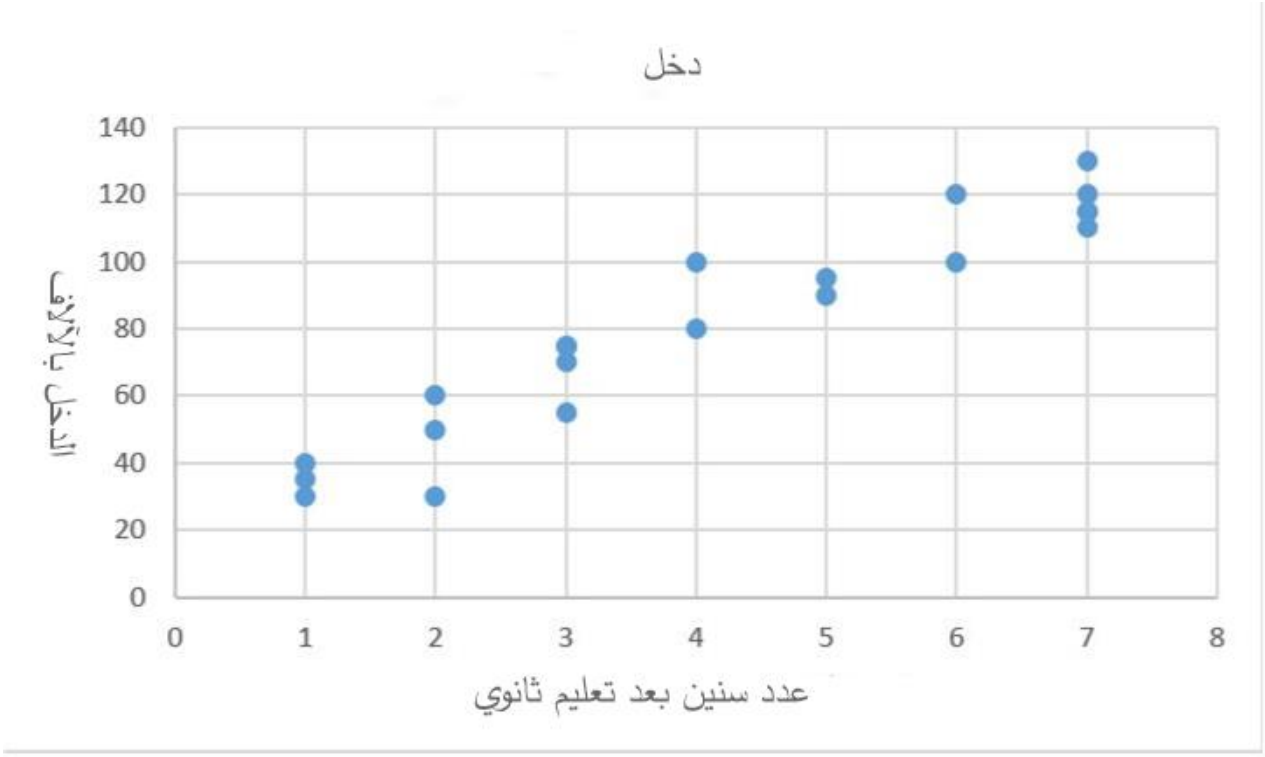
الحسر الخطي (طريقة القيمة الصغرى لمجموع المربعات)

"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 وحدة تـوازي الـف دولار
```



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)

$$ص = ب_0 + ب_1 س + ي$$

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:

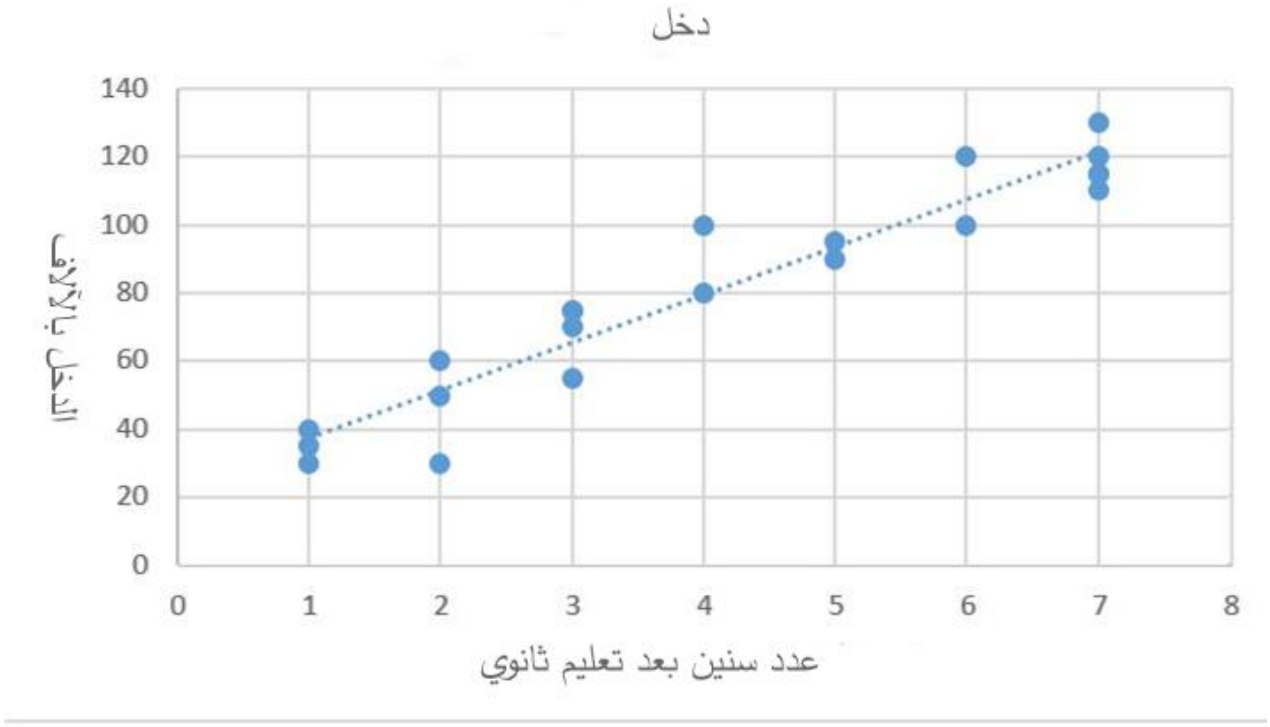
1. قم بتعريف اقتران تكلفة* او الخسارة* , حيث يقوم هذا الاقتران بقياس مدى دقة توقع نموذجك

2. 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*



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

$$\text{التكلفة} = \frac{\sum_{1}^{ن} ((\text{ص}_{\text{ه}} - (\text{ب}_{0} + \text{ب}_{1}\text{س}_{\text{ه}}))^{2}}{2 * \text{ن}}$$

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

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 و TensorFlow 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؟

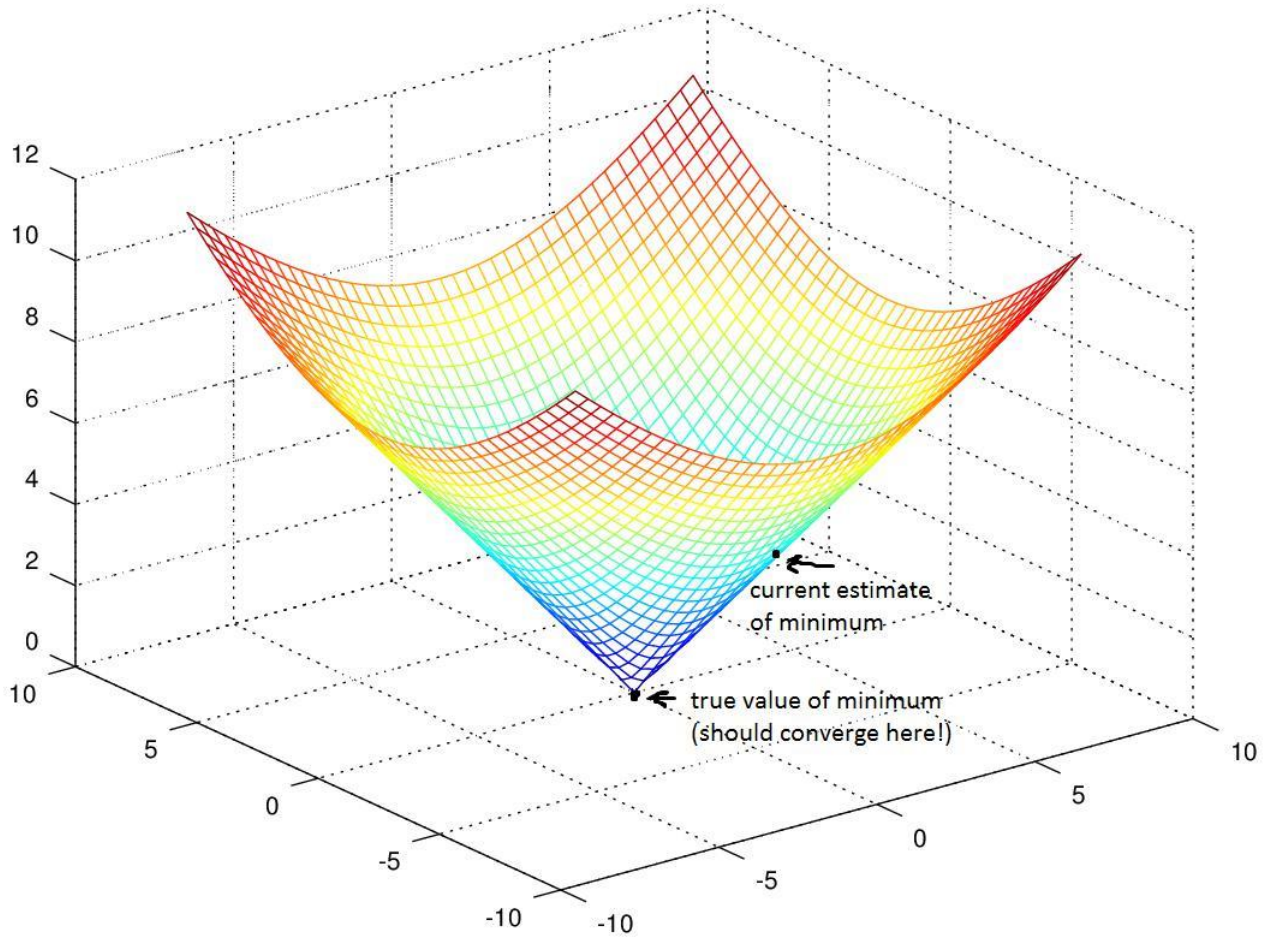
طريقة منطقية ستكون إن تتحس الأرض حولك وتتوجه نحو الأتجاه الذي تنحدر فيه الأرض بشدة . خذ خطوى واعد الفعل حتى تصل إلى ارض مسطحة . وعندها تعرف انك وصلت قاع الواد : إذ تحركت في اي اتجاه ستنتهي في نفس الأرتفاع او اعلى قليلاً

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 :

$$\text{التكلفة} = \frac{\sum_{1}^{ن} ((ص_{ـ} - (ب_{0} + ب_{1} س_{ـ}))^{2}}{2 * ن}$$

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 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

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/ دب 1 هو رقم سلبي ، ثم زيادة ب 1 أمر جيد لأنه سيقلل من الخسارة الكلية. إذا إنه رقم موجب ، فأنت تريد تقليل ب .1إذا كان صفرًا ، فلا تغير ب 1لأنها 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 " :(٨) وهذه القوة التي سألكمك بها"

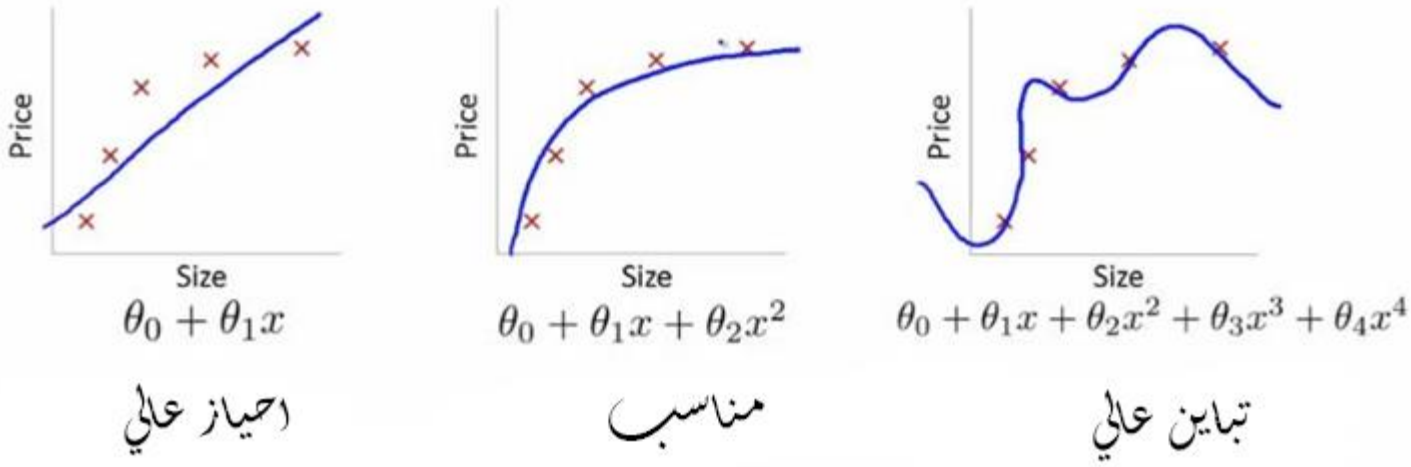
هناك مشكلة شائعة في تعلم الآلة وهي الأكداء : تعلم نموذج يشرح بيانات تدريب بشكل مطابق . ولكن لا يعمم على بيانات الأختبار . يحدث الأكداء عندما يتم يتعلم نموذج بشكل زائد من بيانات تدريب لدرجة إنه يجد خصويات لا تمثل الأنماط في العالم الواقعي . هذا يمثل مشكلة خصوصاً عندما يزيد تعقيد نموذجك . الأرخاء يحدث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

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.

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 .*

$$\text{التكلفة} = \sum_{1}^{ن} \frac{((ص_{م} - (ب_{0} + ب_{1} س_{م})))^{2}}{2 * ن} + \lambda \sum_{هـ=0}^{1} ب_{هـ}^{2}$$

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 !

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) باستخدامPython 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.

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 المُقرض 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 شكلS الذي سترونه في وقت لاحق.

$$\text{س}(س) = \frac{1}{1 + \text{هـ}^{-س}}$$

اقتران دالة سينية 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:

$$\text{غ}(س) = \text{بـ}_0 + \text{بـ}_1 س + \text{ي}$$

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:

$$\text{ح}(\text{ص} = 1) = \text{ق}(\text{غ}(س)) = \frac{1}{1 + \text{هـ}^{-(\text{بـ}0 + \text{بـ}1 س + \text{ي})}}$$

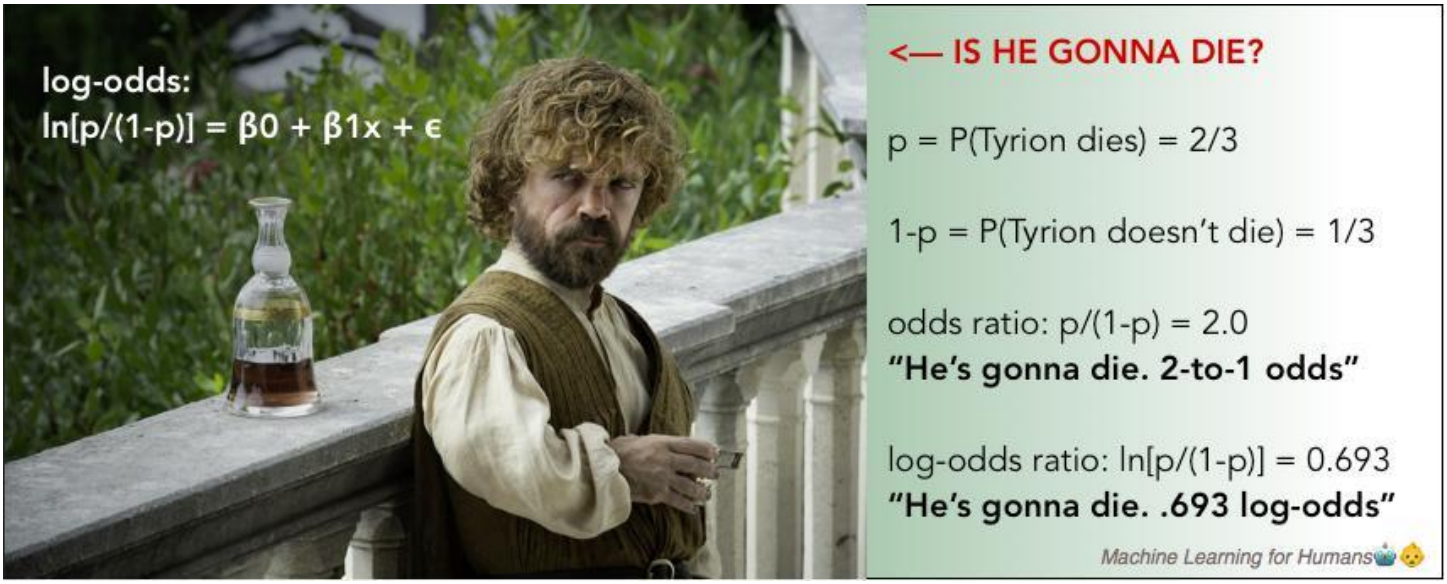
بمعنى آخر ، نحن نحسب احتمال أن يكون مثال التدريب ينتمي إلى فئة معينة.(P (Y = 1 :

هنا قمنا بمعزل p ، احتمال أنY = 1 ، على الجانب الأيسر للمعادلة. إذا كنا نرغب في حل للحصول على نظيف نظيف $\beta 0 + \beta 1 x + \epsilon$ 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

$$\text{لو}\left(\frac{\text{ح}}{1 - \text{ح}}\right) = \text{بـ}_0 + \text{بـ}_1 س + \text{ي}$$

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

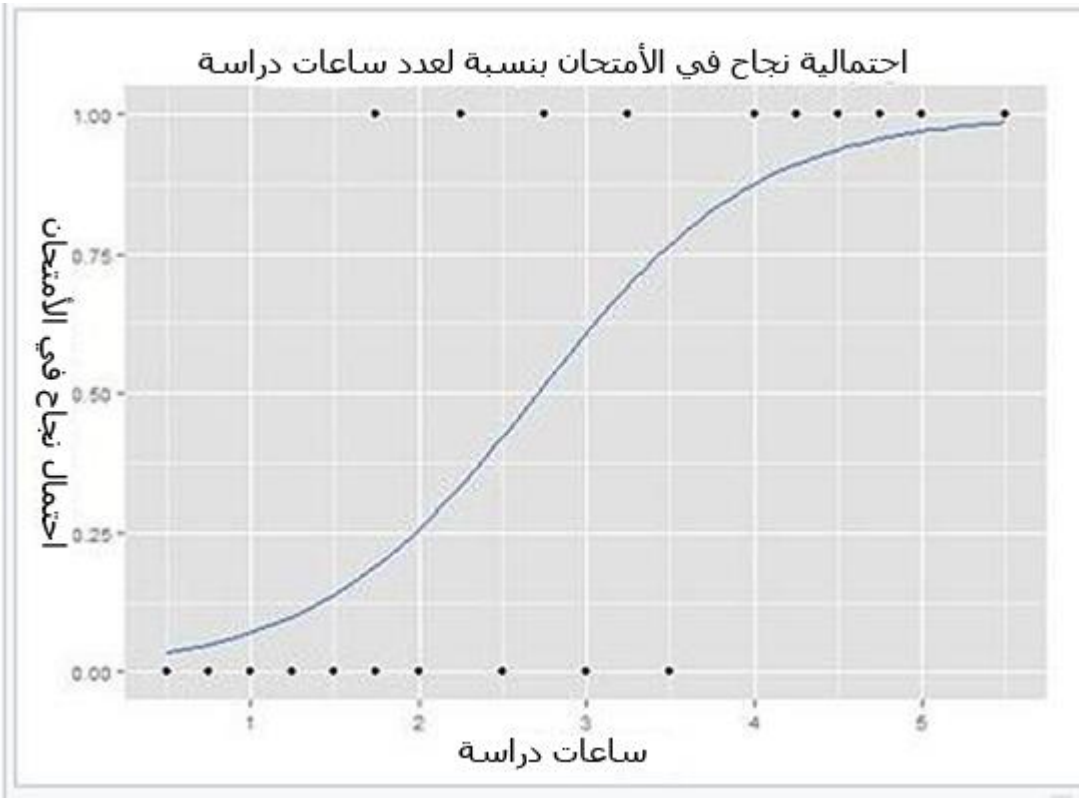
> *"هآي , ما هو احتمال إن تيرون لانستر سيموت في هذا الموسم ""*
> *(كارثة الموسم ثامن جعلت هذا المرجع موجعا ☹)*
>
> *هم. من المؤكد أنه من المرجح أن يحدث مرتين أكثر من ذلك. 2 إلى 1 خلاف. بالتأكيد ، قد "*
> *". ..يبدو من الأهمية بمكان أن يُقتل ، لكننا رأينا جميعاً ما فعلوه مع نيد ستارك*



لاحظ أنه في نموذج اللوغاريثمي، يمثل b_1 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.

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

.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.

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.

$$Cost(\beta) = \frac{\sum_{i=1}^{n}(y^i log(h_\beta(x^i)) + (1-y^i) log(1-h_\beta(x^i))^2}{2*n} + \lambda \sum_{j=1}^{k} \beta_j^2$$

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

تأيد متجهة الآلة (امل)

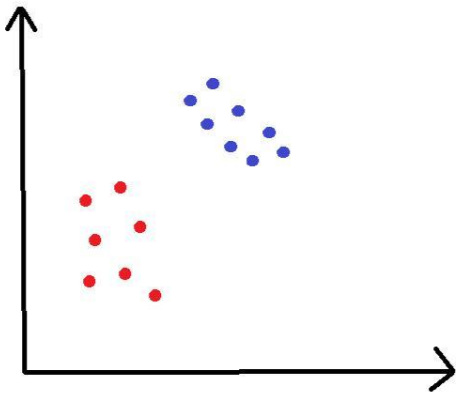
"نحن في غرفة مليئة بكرات زجاجية مرة أخرى. لماذا نحن دائما في غرفة مليئة بكرات زجاجية ؟ كان من الممكن أن أقسم أننا فقدناهم بالفعل."

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.



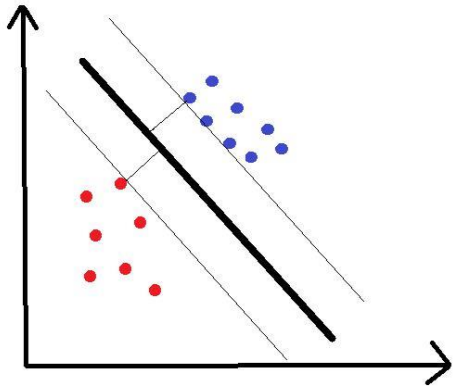
في هذا المثال ، لدينا نقاط في فضاء ثنائية الأبعاد إما حمراء أو زرقاء ، ونود أن نفصل بين الاثنين.

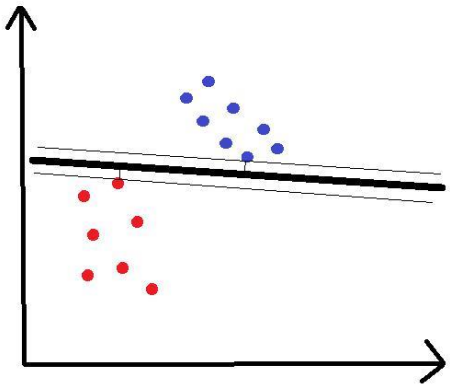
بيانات تدريب تم رسمها . نود إن نصنف نقطة جديدة غير مصنفة . لفعل ذالك تقوم (أمل) برسم خط فاصل (وفي حال كان لدينا فضاء ثلاثي الابعاد ترسم مسطح) لفصل الفضاء إلى منطقة حمراء و منطقة زرقاء . تستطيع تخيل الخط الفاصل الذي سيفصل المجموعتان.

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.

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 لتحقيق هذا الحد الأقصى ، ولكن التفاصيل خارج نطاقنا. لاستكشافها بشكل أكبر ، إليك محاضرة 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.

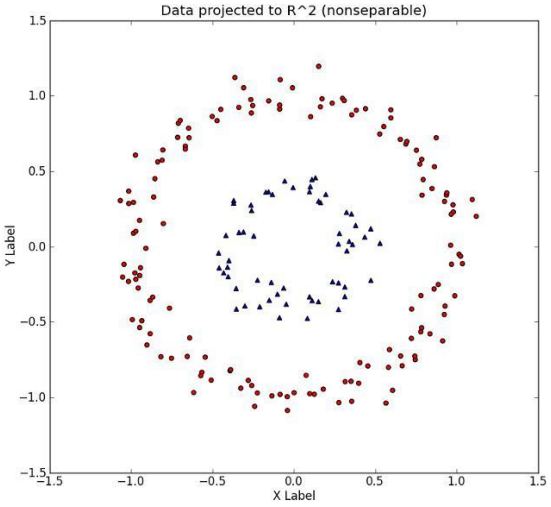
2.ماذا يحدث إذا لم تتمكن من فصل البيانات بشكل كامل؟

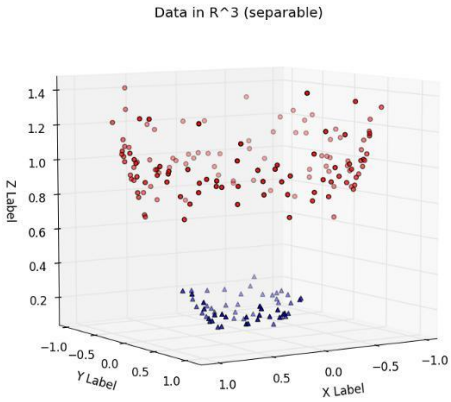
يوجد منهجان لتعمل مع هذه المشكلة

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.

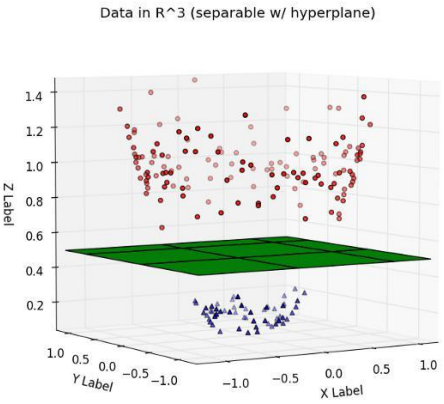
2.2 نقوم برفع نقاط إلى ابعاد عليا
مصنفات غير خطية عن طريق زيادة عدد الأبعاد ، على سبيل المثال تضمينx^2 ، x^3 ، حتى cos (x) ، إلخ. فجأة ، يكون لديك حدود يمكن أن تبدو متعرجة أكثر عندما نعيدها إلى التمثيل الأقل بعدًا.

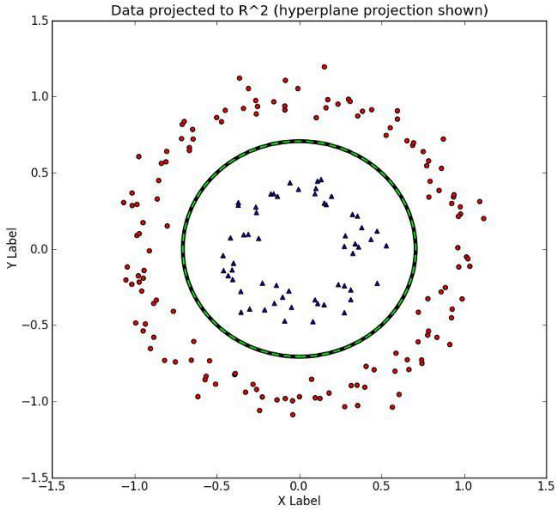
بشكل بدهي ، هذا يشبه وجود كرات أحمر وأزرق ملقى على الأرض بحيث لا يمكن فصله بشكل نظيف عن طريق خط - ولكن إذا تمكنت من جعل كل الرخام الأحمر يطف عن الأرض بالطريقة الصحيحة تمامًا ، يمكنك رسم مسطح يفصلهما. ثم سمحت لهم بالعودة إلى الأرض لمعرفة أين تتوقف الكرات زرقاء وتبدأ الحمراء.





مجموعة بيانات غير قابلة للفصل في فضاء ثنائي الأبعاد $ر^2$ ، 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) :





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.

للتصنيف مع فئتين. يحاولون العثور على طائرة تفصل بين الفئتين بشكل نظيف. عندما لا يكون (أمل) ، يتم استخدام 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.

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

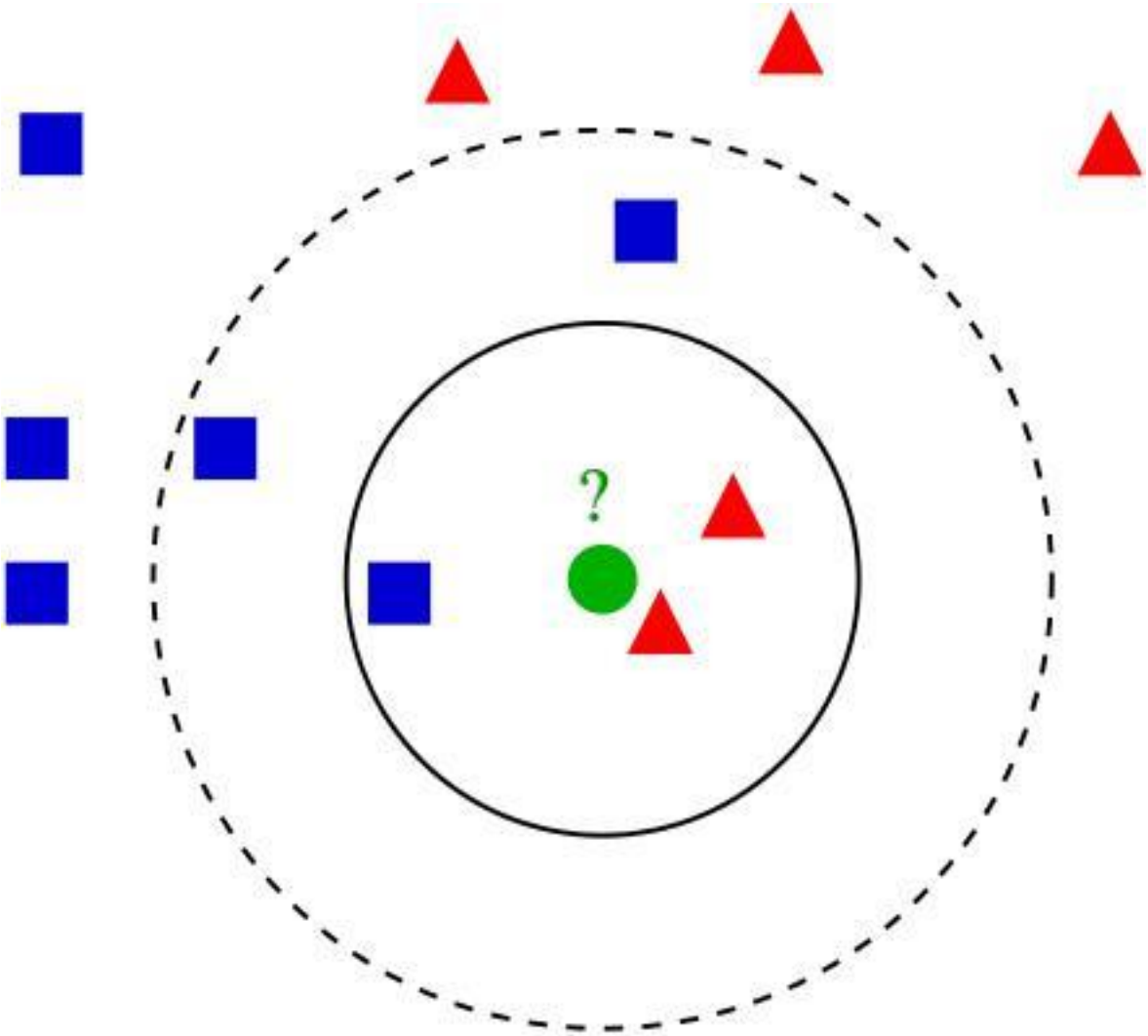
الجار الأقرب لـ(ك) : (جارك)

"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 ؟

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)

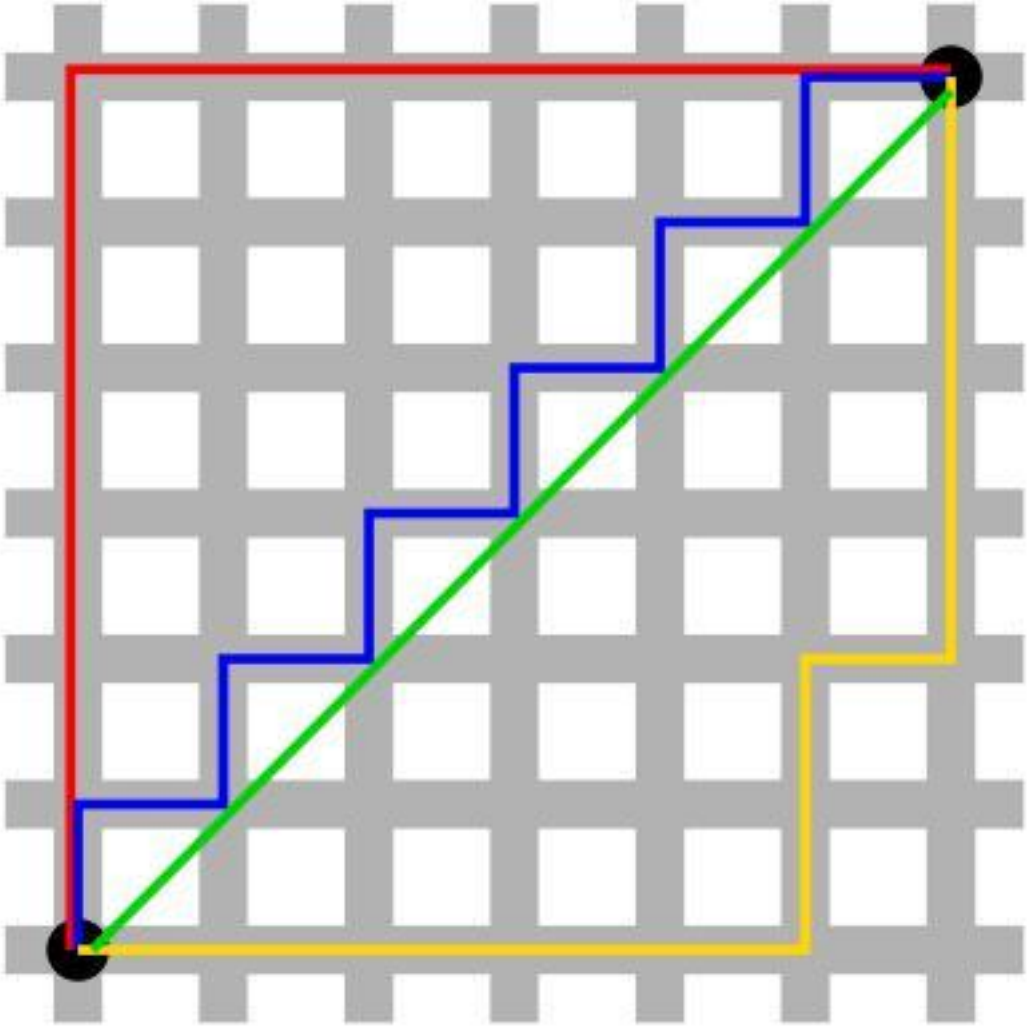
3) خذ الوسيط لأقرب (ك) من المنازل , الوسيط سيكون سعر المنزل

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 ؟

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؟

$$a^2 + b^2 = c^2.$$

c = طول الوتر (الخط الأخضر أعلاه). a و b = طول الضلعين الآخرين بزاوية قائمة (خطوط حمراء أعلاه).

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).

$$c = \sqrt{a^2 + b^2}.$$

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

$$\mathrm{d}(\mathbf{p}, \mathbf{q}) = \mathrm{d}(\mathbf{q}, \mathbf{p}) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \cdots + (q_n - p_n)^2}$$

$$= \sqrt{\sum_{i=1}^{n} (q_i - p_i)^2}.$$

. صيغة للمسافة الإقليدية ، مشتقة من نظرية فيثاغورس.

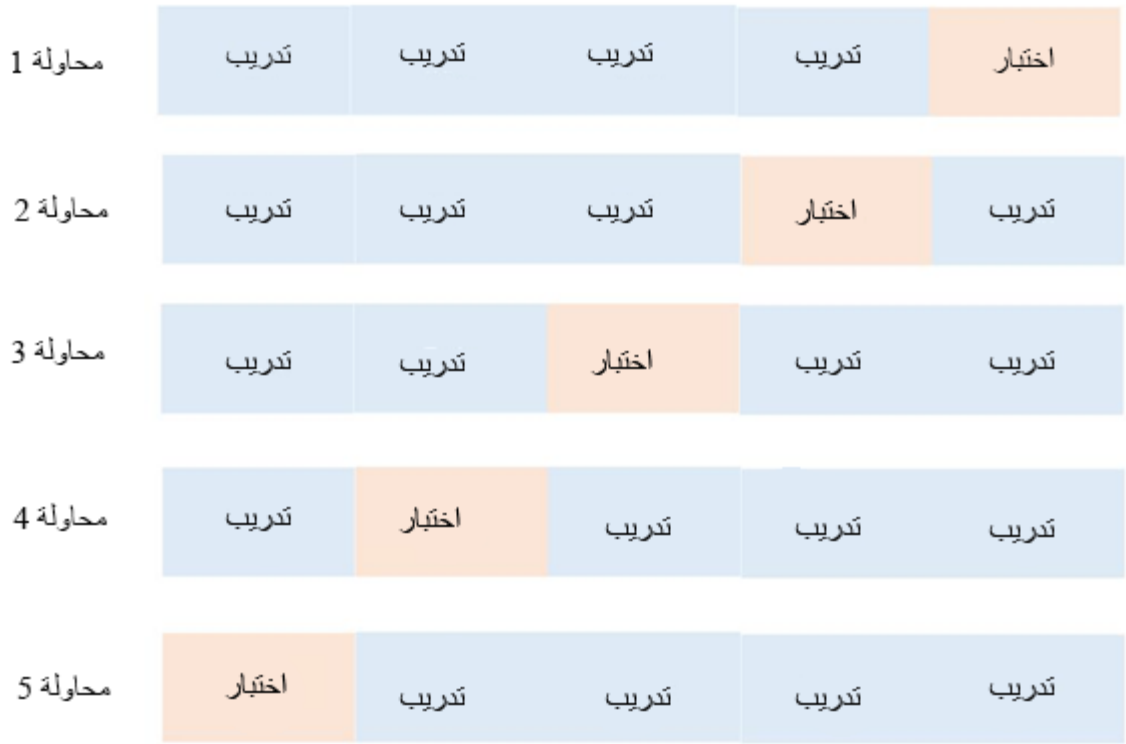
باستخدام هذه الصيغة ، يمكنك حساب قرب جميع نقاط بيانات التدريب من نقطة البيانات التي تحاول تصنيفها ، واتخاذ متوسط / وضع k الأقرب جيرانًا لعمل توقعك.

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

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.



أعلى (**ك**) يمنع الاكداء

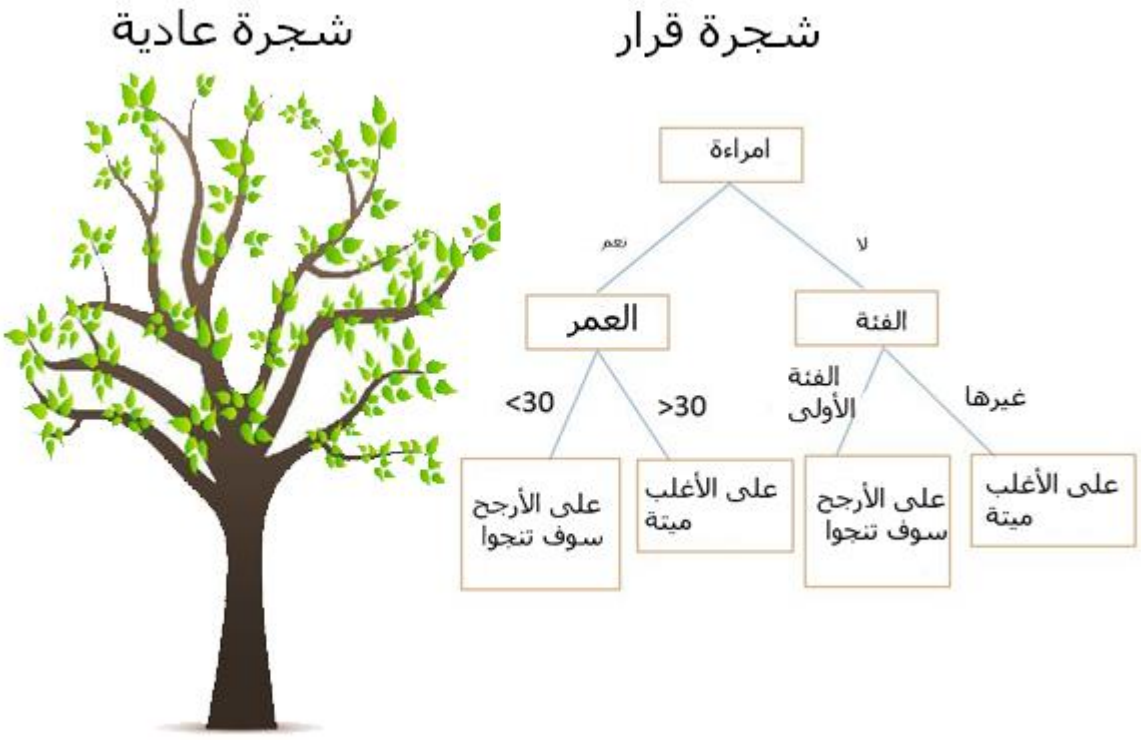
تساعد القيم الأعلى لـ (**ك**) في معالجة التجاوز ، ولكن إذا كانت قيمة k عالية جدًا ، فسيكون نموذجك متحيزًا للغاية وغير مرن. لنأخذ مثالًا متطرفًا: إذا كان (**ن**) = (**ك**) (العدد الإجمالي لنقاط البيانات) ، فإن النموذج سيصنف جميع بيانات الاختبار على أنها متوسط أو وضع بيانات التدريب.

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.

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



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.

اختيار تقسيمات في شجرة القرار

الفوضى هو مقدار الاضطراب في مجموعة (تقاس بمؤشر جيني أو الانتروبيا). 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.

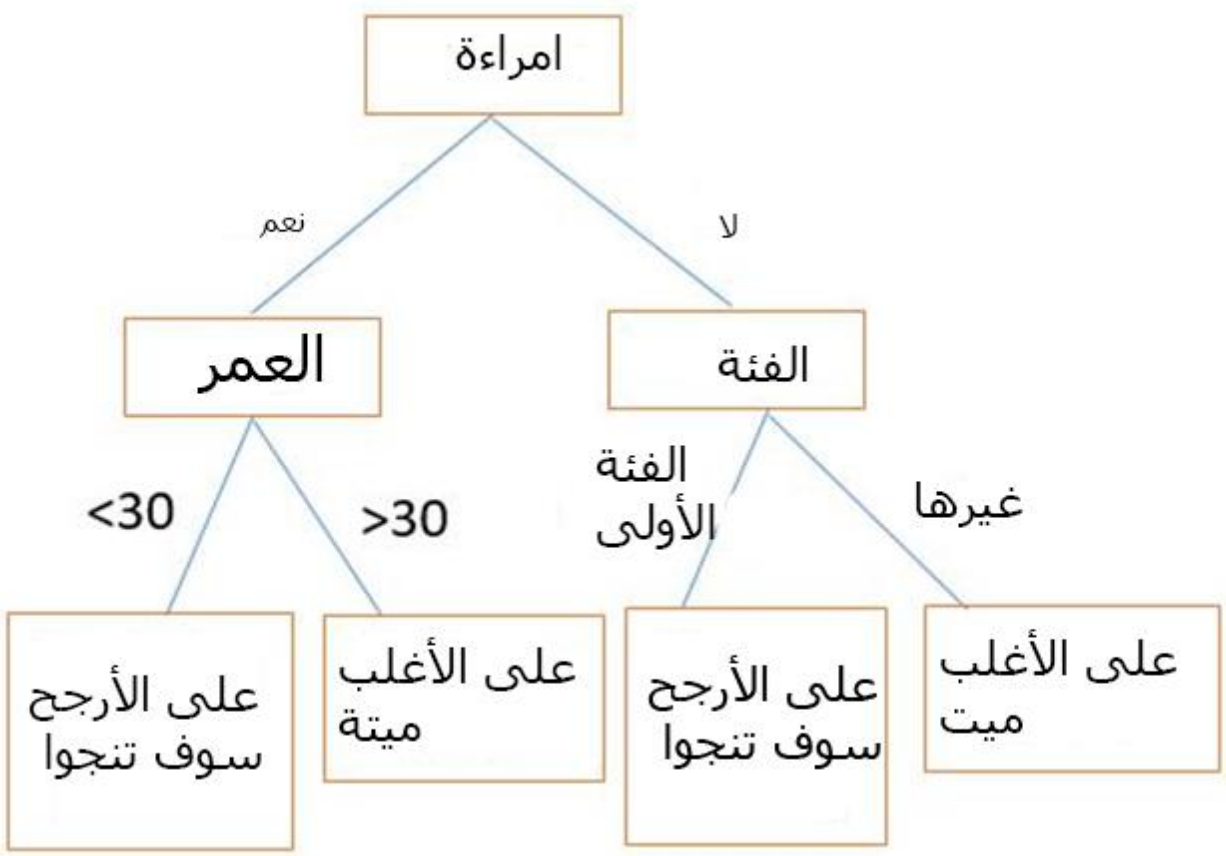
Let's take a look at our Titanic variables:

Data Dictionary

Variable	Definition	Key
survival	Survival	0 = No, 1 = Yes
pclass	Ticket class	1 = 1st, 2 = 2nd, 3 = 3rd
sex	Sex	
Age	Age in years	
sibsp	# of siblings / spouses aboard the Titanic	
parch	# of parents / children aboard the Titanic	
ticket	Ticket number	
fare	Passenger fare	
cabin	Cabin number	
embarked	Port of Embarkation	C = Cherbourg, Q = Queenstown, S = Southampton

Source: Kaggle

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



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.

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.

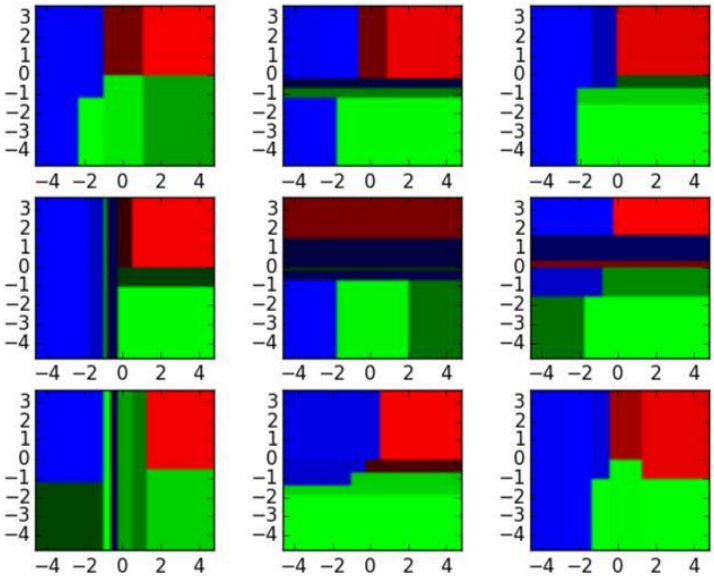
: يمكن لشجرة قرار واحدة القيام بالكثير من الأخطاء لأنها تحتوي على أحكام بالأبيض والأسود. الغابة العشوائية عبارة عن مقدر تقديري يقوم بتجميع العديد من أشجار القرارات ، مع بعض التعديلات المفيدة:

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.

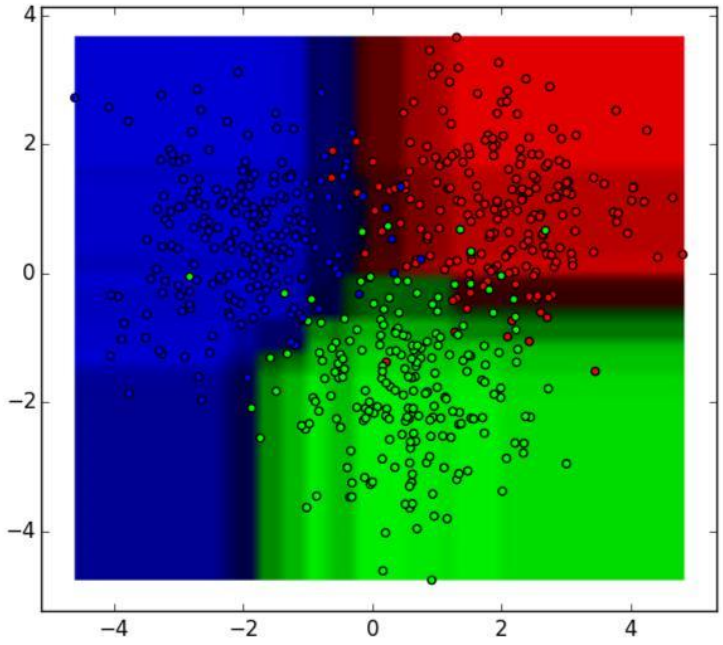
تمنع هذه التعديلات أيضًا الترابط الشديد بين الأشجار. بدون رقم 1 ورقم 2 أعلاه ، ستكون كل شجرة متطابقة ، لأن الانقسام الثنائي التعاودي أمر حتمي.

. للتوضيح ، راجع تصنيفات أشجار القرار التسع هذه أدناه.



Source: http://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/

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.

وبهذا نختتم دراستنا للتعلم المشرف!

احسنت. لقد غطينا في هذا القسم:

- **خورزميتان بدون معاملات لتعلم المشرف (جارك) و أشجار القرار**
- **مقاييس المسافة وكسب المعلومات**
- **الغابات العشوائية كمثال على نموذج تجميعي**
- **تصديق المتقاطع و ضبط المتغيرات الفائقة**

نأمل أن يكون لديك الآن بعض البديهيات القوية لكيفية تعلمنا في ضوء مجموعة بيانات التدريب واستخدامها لعمل تنبؤات باستخدام بيانات الاختبار.

بعد ذلك ، سنتحدث عن كيفية التعامل مع المشكلات التي لا تتوفر لدينا فيها أي بيانات تدريبية مصنفة للعمل معها ، في الجزء 3: التعلم الغير مشرف.

مواد تدريبية وقراءات إضافية
2.3 – تطبيق (جارك)
جرب هذه الإرشادات لتنفيذ (جارك) من البداية في بايثون .قد ترغب أيضًا في إلقاء نظرة على وثائق-scikit Learnللتعرف على كيفية عمل التطبيقات سابقة البناء.

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)

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.

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.

1Y STARTING OUT	**8X LARGE HOUSEHOLDS**	**15M TOP WEALTH**
Cluster 39 Setting Goals	Cluster 11 Schools & Shopping	Cluster 2 Established Elite
Cluster 45 Offices & Entertainment	Cluster 12 On the Go	Cluster 3 Corporate Connected
Cluster 57 Collegiate Crowd	Cluster 19 Country Comfort	
Cluster 58 Outdoor Fervor	Cluster 27 Tenured Proprietors	**16M LIVING WELL**
Cluster 67 First Steps		Cluster 14 Career Centered
	9B COMFORTABLE INDEPENDENCE	Cluster 15 Country Ways
2Y TAKING HOLD	Cluster 29 City Mixers	Cluster 23 Good Neighbors
Cluster 18 Climbing the Ladder	Cluster 35 Working & Active	
Cluster 21 Children First	Cluster 56 Metro Active	**17M BARGAIN HUNTERS**
Cluster 24 Career Building		Cluster 43 Work & Causes
Cluster 30 Out & About	**10B RURAL-METRO MIX**	Cluster 44 Open Houses
	Cluster 47 Rural Parents	Cluster 55 Community Life
3Y SETTLING DOWN	Cluster 53 Metro Strivers	Cluster 63 Staying Home
Cluster 34 Outward Bound	Cluster 60 Rural & Mobile	Cluster 68 Staying Healthy

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.

What's My Cluster ?

Cluster #24: Career Building

Career Building singles are young, but well compensated. While repaying their education loans they are beginning to save and invest. They favor trendy stores that cater to their age range, incomes and aspirations, such as Express, H&M and Sephora. They enjoy new technology, and read magazines on mobile devices. They visit The Apple Store. They use the Internet extensively for entertainment news, music, podcasts and services. Sports are important, too, either as a fan or a participant. They listen to football, watch MMA and have fun skiing and playing volleyball.

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.

دعنا نتعرف على طريقتين للتجميع لتطوير حدس لكيفية تنفيذ هذه المهمة.

العنقدة بالوسطاء المتعددين

"وأعطيت حلقات k لسباق نقاط المركزية ، الذين يرغبون قبل كل شيء في القوة".

الهدف من التجميع هو إنشاء مجموعات من نقاط البيانات بحيث تختلف النقاط في مجموعات مختلفة بينما تتشابه النقاط داخل المجموعة.

مع العنقدة بالوسطاء المتعددين، نريد تجميع نقاط البيانات الخاصة بنا في مجموعات .k يُنشئ k الأكبر مجموعات أصغر بمزيد من التفصيل ، بينما يعني k الأدنى مجموعات أكبر وتفاصيل أقل.

سيكون ناتج الخوارزمية عبارة عن مجموعة من "العلامات" التي تخصص كل نقطة بيانات لإحدى مجموعات .k في العنقدة بالوسطاء المتعددين ، الطريقة التي يتم بها تعريف هذه المجموعات هي عن طريق إنشاء النقطه الوسطى لكل مجموعة. النقط الوسطى هي مثل قلب الكتلة ، فهي "تلتقط" النقاط الأقرب إليها وتضيفها إلى الكتلة.

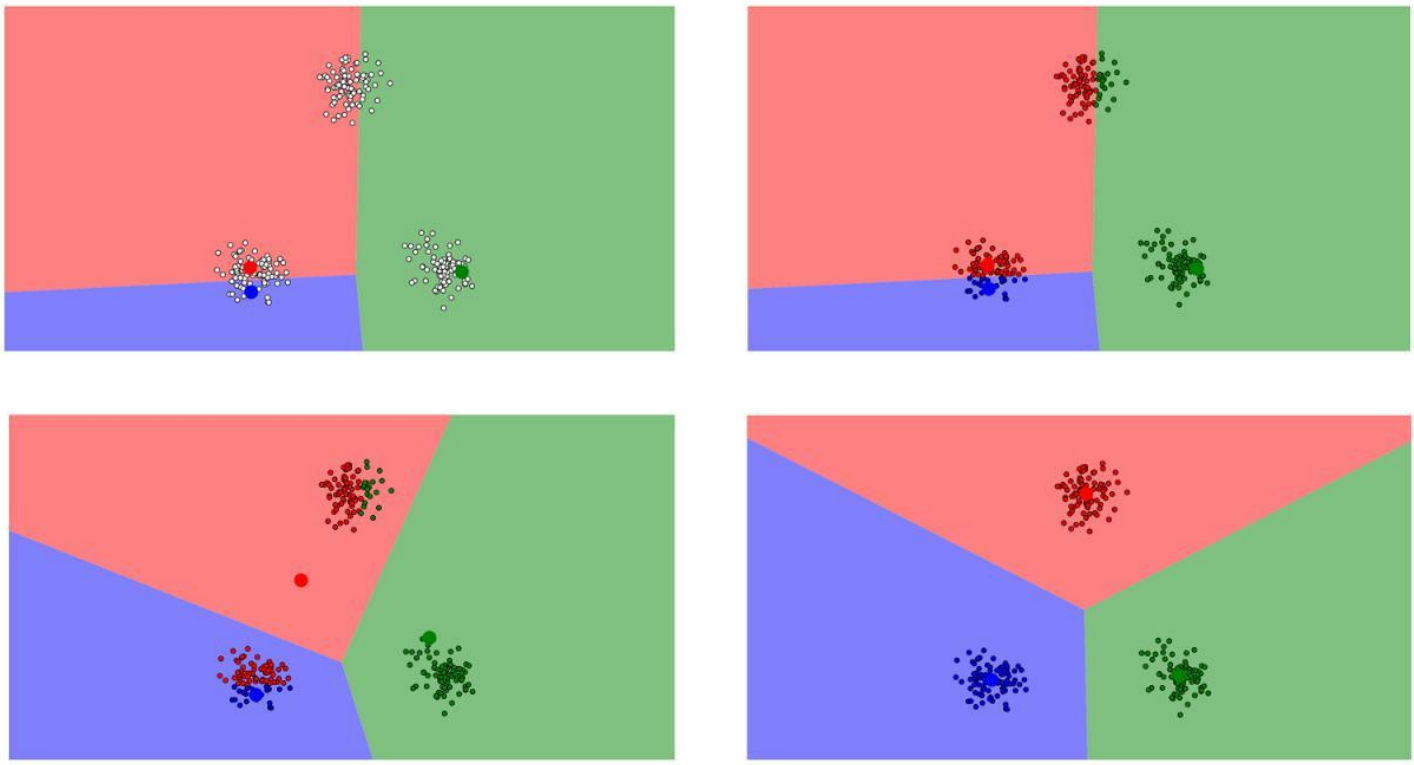
فكر في هؤلاء على أنهم الأشخاص الذين يظهرون في حفلة وسرعان ما يصبحون مركز الاهتمام لأنهم يتمتعون بجاذبية كبيرة. إذا كان هناك واحد منهم فقط ، فسيتجمع الجميع ؛ إذا كان هناك الكثير ، فسيتم تشكيل العديد من مراكز النشاط الأصغر.

: فيما يلي خطوات العنقدة بالوسطاء المتعددين:

1. تحديد النقط الوسطى ك. قم بتهيئة هذه بشكل عشوائي (هناك أيضًا خوارزميات مربي الحيوانات لتهيئة النقط الوسطى التي ينتهي بها الأمر إلى التقارب بشكل أكثر فعالية).
2. البحث عن أقرب النقطه الوسطى وتحديث تعيينات الكتلة. قم بتعيين كل نقطة بيانات لإحدى مجموعات .k يتم تعيين كل نقطة بيانات لأقرب مجموعة من النقط الوسطى. هنا ، مقياس "القرب" هو معلمة مفرطة - غالبًا مسافة إقليدية.
3. انقل النقط الوسطى إلى مركز مجموعاتهم. يتم حساب الموضع الجديد لكل النقطه الوسطى كمتوسط موضع لجميع النقاط في مجموعتها.

استمر في تكرار الخطوتين 2 و 3 حتى تتوقف النقطه الوسطى عن التحرك كثيرًا في كل تكرار (على سبيل المثال ، حتى تتقارب الخوارزمية)

هذا ، باختصار ، كيف يعمل التجميع العنقدة بالوسطاء المتعددين !تحقق من هذا التصور للخوارزمية - اقرأه مثل كتاب فكاهي. يتم تلوين كل نقطة في الطائرة وفقًا للنقطة الوسطى الأقرب إليها في كل لحظة. ستلاحظ أن النقط الوسطى (الدوائر الكبيرة ذات اللون الأزرق والأحمر والأخضر) تبدأ بشكل عشوائي ثم تتكيف بسرعة لالتقاط مجموعاتها الخاصة.



من التطبيقات الواقعية الأخرى (خوارزمية تصنيفية) تصنيف الأرقام المكتوبة بخط اليد. لنفترض أن لدينا صورًا للأرقام كمتجه طويل لسطوع البكسل. لنفترض أن الصور بالأبيض والأسود هي 64 × 64 بكسل. يمثل كل بكسل بعدًا. لذا فإن العالم الذي تعيش فيه هذه الصور له أبعاد 64 × 64 = 4096. في هذا العالم المكون من 4096 بعدًا ، تتيح لنا (خوارزمية تصنيفية) تجميع الصور القريبة من بعضها ونفترض أنها تمثل نفس الرقم ، مما يمكن أن يحقق نتائج جيدة جدًا للتعرف على الأرقام.

تجميع هرمي

"دعونا نجعل مليون خيار لتصبح سبعة خيارات. أو خمسة. أم عشرين؟ مه ، يمكننا أن نقرر لاحقًا."

يشبه التجميع الهرمي التجميع المنتظم ، باستثناء أنك تهدف إلى بناء تسلسل هرمي من المجموعات. يمكن أن يكون هذا مفيدًا عندما تريد المرونة في عدد المجموعات التي تريدها في النهاية. على سبيل المثال ، تخيل تجميع العناصر في في الصفحة الرئيسية ، قد ترغب في بعض الفئات العامة للعناصر .Amazon أو Etsy سوق عبر الإنترنت مثل الخاصة بها

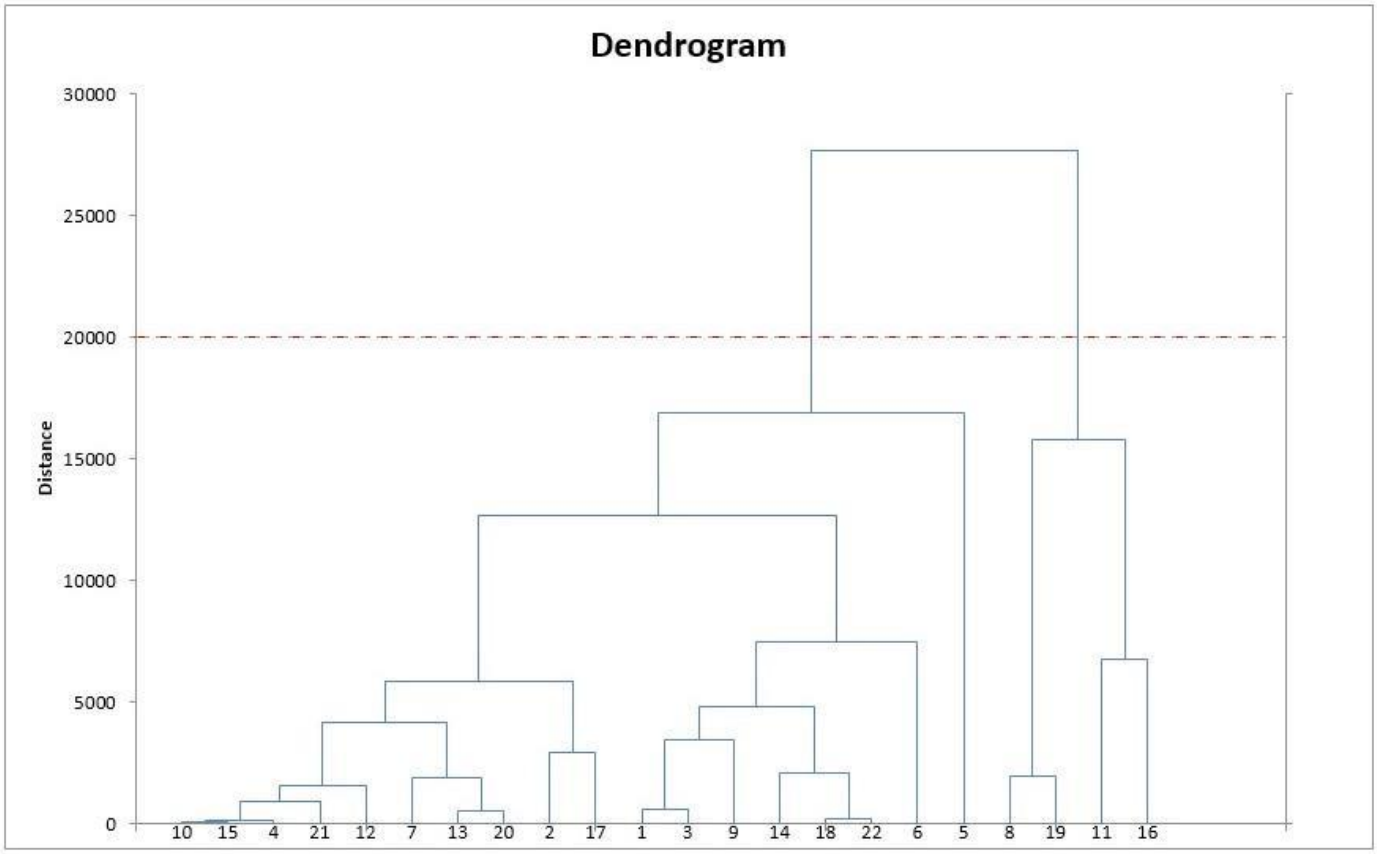
تنقل بسيط ، ولكن كلما انتقلت إلى فئات تسوق أكثر تحديدًا ، قد ترغب في زيادة مستويات الدقة ، أي مجموعات أكثر تميزًا من العناصر.

فيما يتعلق بالمخرجات من الخوارزمية ، بالإضافة إلى تعيينات المجموعة ، يمكنك أيضًا بناء شجرة لطيفة تخبرك عن التسلسلات الهرمية بين المجموعات. يمكنك بعد ذلك اختيار عدد المجموعات التي تريدها من هذه الشجرة.

فيما يلي خطوات التجميع الهرمي

1. ابدأ بمجموعات N ، واحدة لكل نقطة بيانات.
2. ادمج المجموعتين الأقرب لبعضهما البعض. الآن لديك مجموعات N-1.
3. أعد حساب المسافات بين المجموعات. هناك عدة طرق للقيام بذلك (راجع هذا البرنامج التعليمي لمزيد من التفاصيل). أحدها (يسمى متوسط الارتباط العنقودي) هو اعتبار المسافة بين مجموعتين متوسط المسافة بين جميع أعضائها.
4. كرر الخطوتين 2 و 3 حتى تحصل على مجموعة واحدة من نقاط البيانات N. تحصل على شجرة تُعرف أيضًا باسم(مخطط شجر الأسنان) مثل الشجرة أدناه.
5. **اختر عددًا من العناقيد وارسم خطًا أفقيًا في مخطط الأسنان.**

على سبيل المثال ، إذا كنت تريد k = 2 عناقيد ، يجب أن ترسم خطًا أفقيًا حول "المسافة = 20000." ستحصل على مجموعة واحدة بها نقاط بيانات 8 و 9 و 11 و 16 ومجموعة واحدة مع بقية نقاط البيانات. بشكل عام ، عدد المجموعات التي تحصل عليها هو عدد نقاط التقاطع لخطك الأفقي مع الخطوط العمودية في مخطط الأسنان.



Source: Solver.com. For more detail on hierarchical clustering, you can check this video out.

تخفيض الأبعاد

"إنها ليست الزيادة اليومية ، بل الزيادة اليومية. تخلص من الأشياء غير الضرورية". - بروسلي

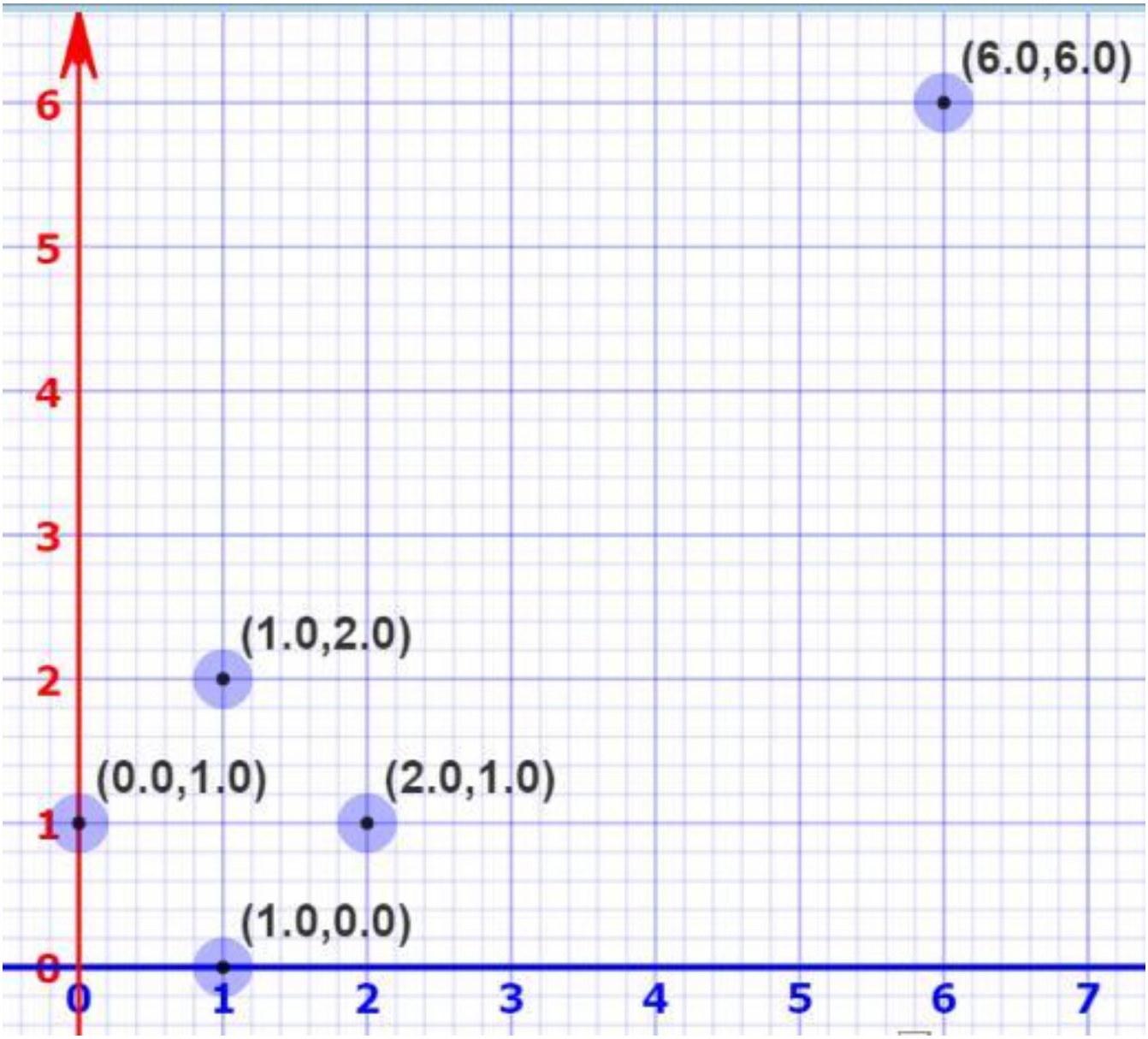
يشبه تخفيض الأبعاد إلى حد كبير الضغط. يتعلق الأمر بمحاولة تقليل تعقيد البيانات مع الاحتفاظ بأكبر قدر ممكن من البنية ذات الصلة. إذا التقطت صورة بسيطة بحجم 128 × 128 × 3 بكسل) الطول × العرض × قيمة (RGB ، فهذا يعني 49152 بعدًا من البيانات. إذا كنت قادرًا على تقليل أبعاد المساحة التي تعيش فيها هذه الصور دون تدمير الكثير من المحتوى ذي المعنى في الصور ، فقد قمت بعمل جيد في تقليل الأبعاد.

. سنلقي نظرة على طريقتين شائعتين في الممارسة: تحليل العناصر الرئيسية (تار) و مجزئ القيمة المفردة (جيم)

تحليل العناصر الرئيسية (تار)

أولاً ، القليل من التنشيط للجبر الخطي - لنتحدث عن المساحات والقواعد.

أنت على دراية بالمستوى الإحداثي ذي الأصل0) O ، (0ومتجهات الأساس1) i ، (0و0) j ، .(1اتضح أنه يمكنك اختيار أساس مختلف تمامًا ولا يزال لديك كل الرياضيات. على سبيل المثال ، يمكنك الاحتفاظ بـ O كأصل واختيار الأساس للمتجهين2) =' i ، (1و1) = j' ، .(2إذا كان لديك الصبر لذلك ، فستقنع نفسك أن النقطة المسماة (2،2) في نظام الإحداثيات i ، j، مسماة (6 ، 6) في نظام i ، .j



Plotted using Mathisfun's "Interactive Cartesian Coordinates"

هذا يعني أنه يمكننا تغيير أساس الفضاء. تخيل الآن مساحة ذات أبعاد أعلى بكثير. مثل ، أبعاد 50 .Kيمكنك تحديد أساس لتلك المساحة ، ثم تحديد 200 متجه مهم فقط لهذا الأساس. تسمى متجهات الأساس هذه بالمكونات الرئيسية ، وتشكل المجموعة الفرعية التي تحددها مساحة جديدة أصغر في الأبعاد من المساحة الأصلية ولكنها تحافظ على أكبر قدر ممكن من تعقيد البيانات.

لتحديد أهم المكونات الأساسية ، ننظر إلى مقدار تباين البيانات الذي تلتقطه وترتيبهم حسب هذا المقياس.

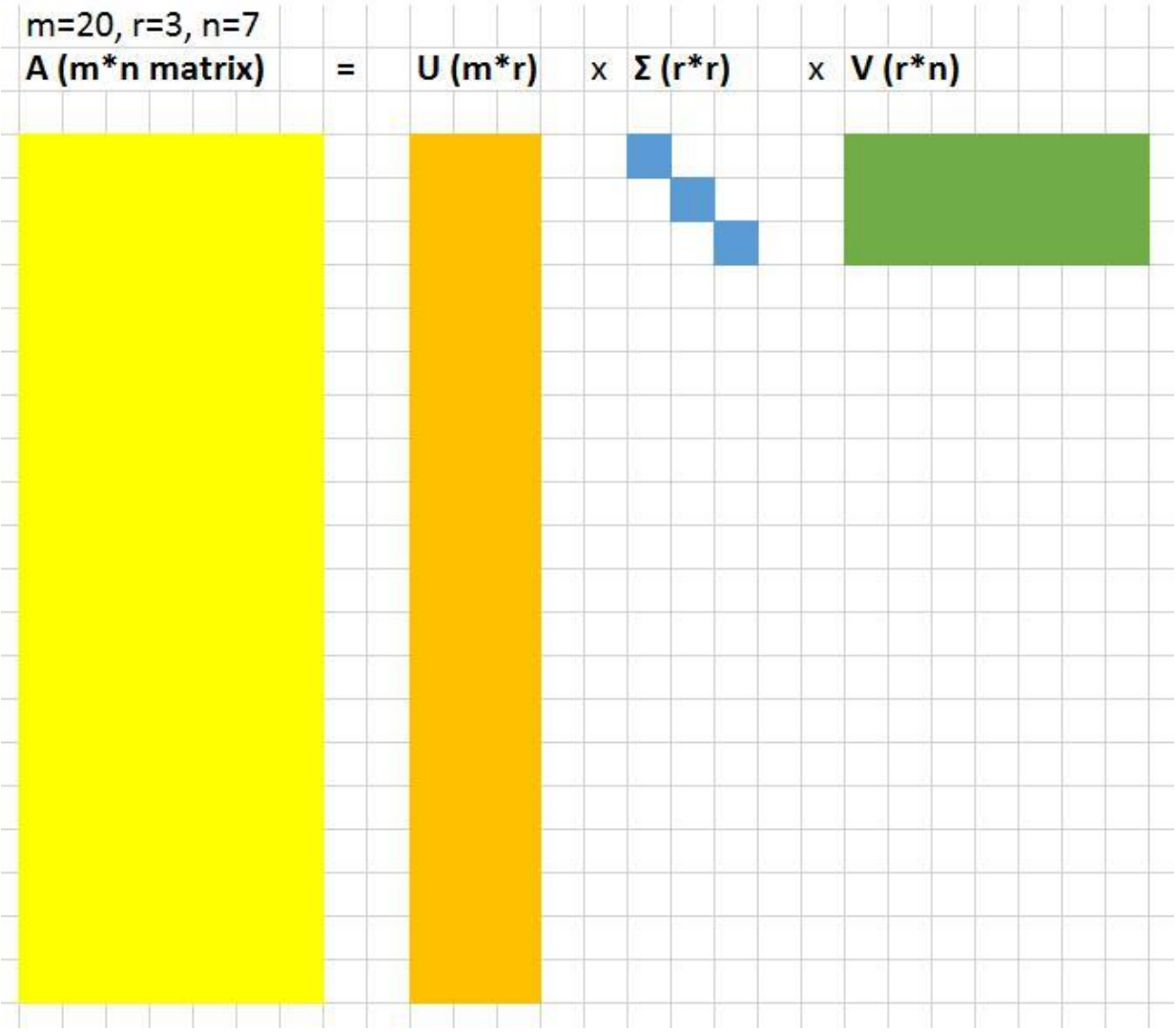
طريقة أخرى للتفكير في هذا الأمر هي أن (تار) يعيد رسم المساحة التي توجد بها بياناتنا لجعلها أكثر قابلية للضغط. البعد المحول أصغر من البعد الأصلي.

من خلال الاستفادة من الأبعاد العديدة الأولى للمساحة المعاد تعيينها فقط ، يمكننا البدء في فهم تنظيم مجموعة البيانات. هذا وعد بتقليل الأبعاد: تقليل التعقيد (الأبعاد في هذه الحالة) مع الحفاظ على الهيكل (التباين). إليك ورقة ممتعة كتبها سامر عن استخدام (تار) (ورسم خرائط الانتشار ، وهي تقنية أخرى) لمحاولة فهم إصدار كابل ويكيليكس.

مجزئ القيمة المفردة (جيم)

دعنا نمثل بياناتنا مثل مصفوفة كبيرةSVD .A = m x n جيم هو حساب يسمح لنا بتحليل تلك المصفوفة الكبيرة إلى منتج من 3 مصفوفات أصغر (U = m x r) ، مصفوفة قطرية Σ = r x r ، و V = r x n حيث r عدد صغير.(

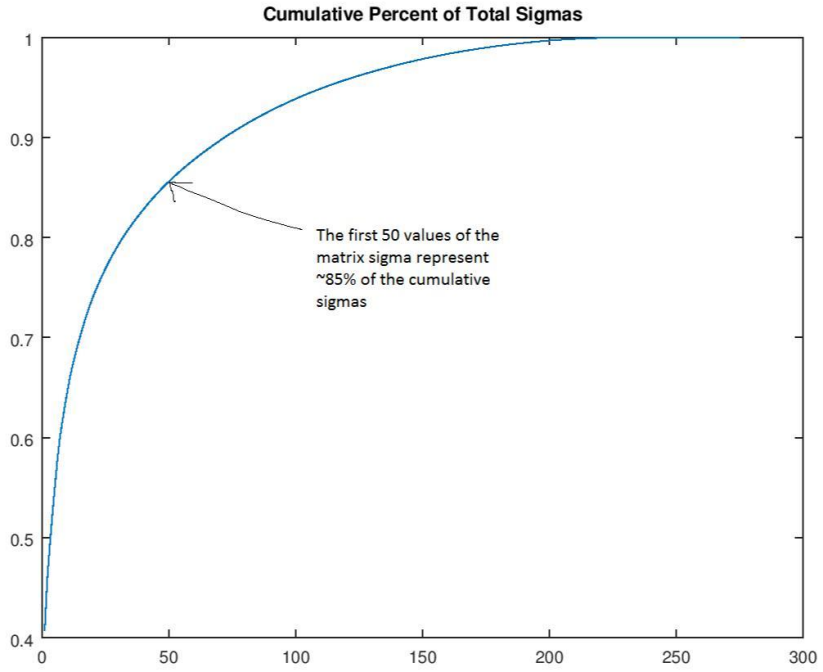
إليك المزيد من التوضيح المرئي لهذا المنتج لتبدأ به:



تسمى القيم الموجودة في المصفوفة r * r القطرية Σ القيم المفردة. ما يميزهم هو أنه يمكن استخدام هذه القيم المفردة لضغط المصفوفة الأصلية. إذا قمت بإسقاط أصغر 20٪ من القيم المفردة والأعمدة المرتبطة بها في المصفوفتين U و V ، فإنك توفر مساحة كبيرة جدًا ولا تزال تحصل على تمثيل لائق للمصفوفة الأساسية.

لفحص ما يعنيه ذلك بشكل أكثر دقة ، دعنا نعمل مع هذه الصورة للكلب:

سنستخدم الشفرة المكتوبة في مشاركة أندرو جيبانسكي على (فيم) .أولاً ، نوضح أنه إذا قمنا بترتيب القيم الفردية) قيم المصفوفة (Σ من حيث الحجم ، فإن أول 50 قيمة فردية تحتوي على 85٪ من حجم المصفوفة بأكملها.Σ



يمكننا استخدام هذه الحقيقة لتجاهل الـ 250 قيمة التالية لـ sigma (أي ضبطها على 0) والاحتفاظ فقط بإصدار "رتبة 50" لصورة الكلب. هنا ، نقوم بإنشاء كلب رتبة 200 ، و 100 ، و 50 ، و 30 ، و 20 ، و 10 ، و 3 كلاب. من الواضح أن الصورة أصغر ، لكن دعنا نتفق على أن الكلب من المرتبة 30 لا يزال جيدًا. الآن دعونا نرى مقدار الضغط الذي نحققه مع هذا الكلب. مصفوفة الصورة الأصلية 305 * 275 = 83875 قيمة. الكلب المصنف 30 هو 305 * 30 + 30 + 30 * 275 = 17430 - قيم أقل بخمس مرات تقريبًا مع خسارة قليلة جدًا في جودة الصورة. سبب الحساب أعلاه هو أننا نتجاهل أيضًا أجزاء المصفوفة U و V التي يتم ضربها بالأصفار عند تنفيذ العملية) UΣ'V حيث Σ' هي النسخة المعدلة من Σ التي تحتوي فقط على أول 30 قيمة فيه.(

غالبًا ما يستخدم التعلم غير الخاضع للإشراف لمعالجة البيانات مسبقًا. عادةً ما يعني ذلك ضغطه بطريقة تحافظ على المعنى كما هو الحال مع (تار) أو (فيم) قبل إطعامه إلى شبكة عصبية عميقة أو خوارزمية تعلم أخرى خاضعة للإشراف.

للأمام!

الآن بعد أن انتهيت من هذا القسم ، حصلت على نكتة مروعة ، مروعة ، لن يتم ذكرها أبدًا مرة أخرى حول التعلم غير الخاضع للإشراف. هنا يذهب...

> *Person-in-joke-#1: Y would u ever need to use unsupervised tho?*
> *Person-in-joke-#2: Y? there's no Y.*

التالي ... الجزء 4: الشبكات العصبية والتعلم العميق!

مواد تدريبية وقراءات إضافية

3a—k-means clustering
Play around with this clustering visualization to build intuition for how the algorithm works. Then, take a look at this implementation of k-means clustering for handwritten digits and the associated tutorial.

3b—SVD
For a good reference on SVD, go no further than Andrew Gibiansky's post.

الجزء رابع : شبكات العصبية و تعلم العميق

أين ، لماذا ، أين ، وكيف تعمل الشبكات العصبية العميقة. وحي من الدماغ. الشبكات العصبية التلافيفية (شبت) والشبكات العصبية المتكررة(شبر) 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:

ص = ق(س)+ي

التدريب: الآلة تتعلم ق من بيانات التدريب المسمى

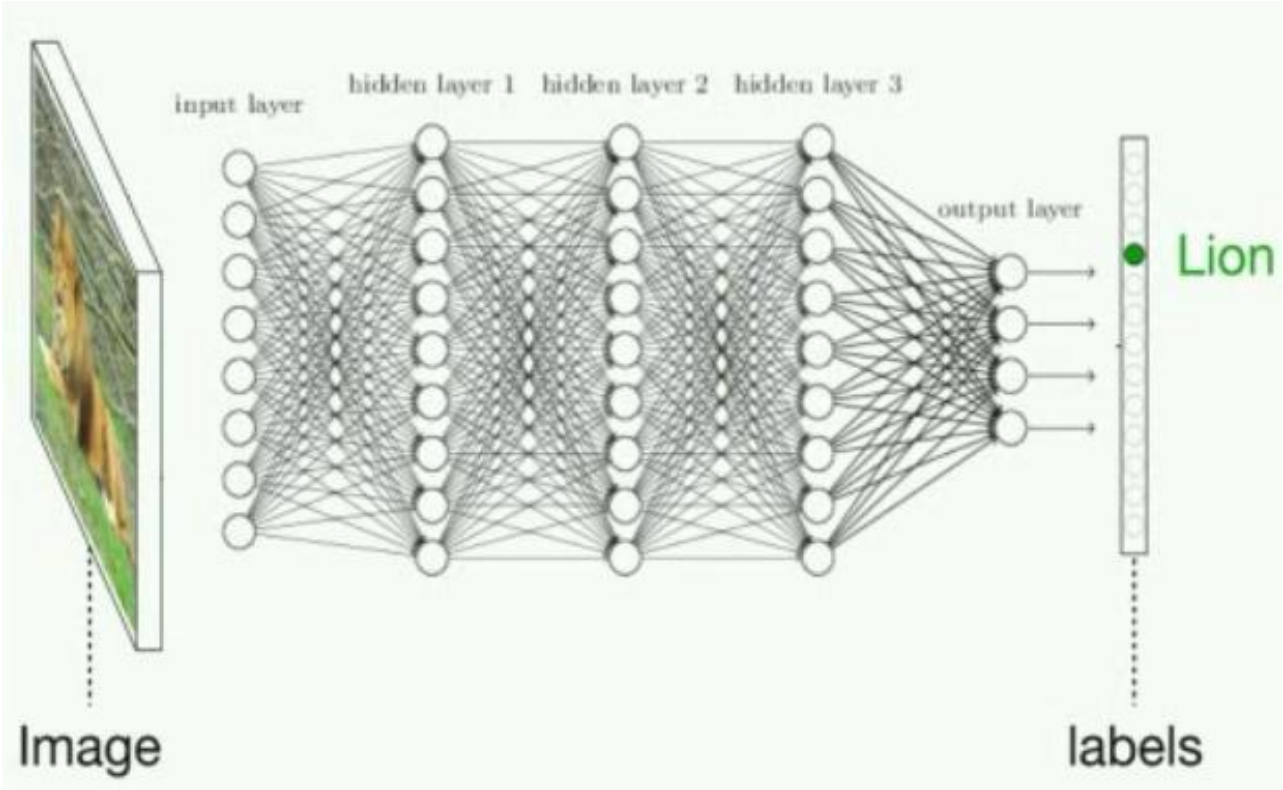
الاختبار: تتوقع الآلة ص من بيانات الاختبار غير المسماة

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 ، 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).

بشكل صوري شبكة عصبية عميقة لتصنيف صور تبدو بهذا شكل:



صورة من نظرة عامة على التعلم العميق لمدة ساعة لجيف كلون

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."

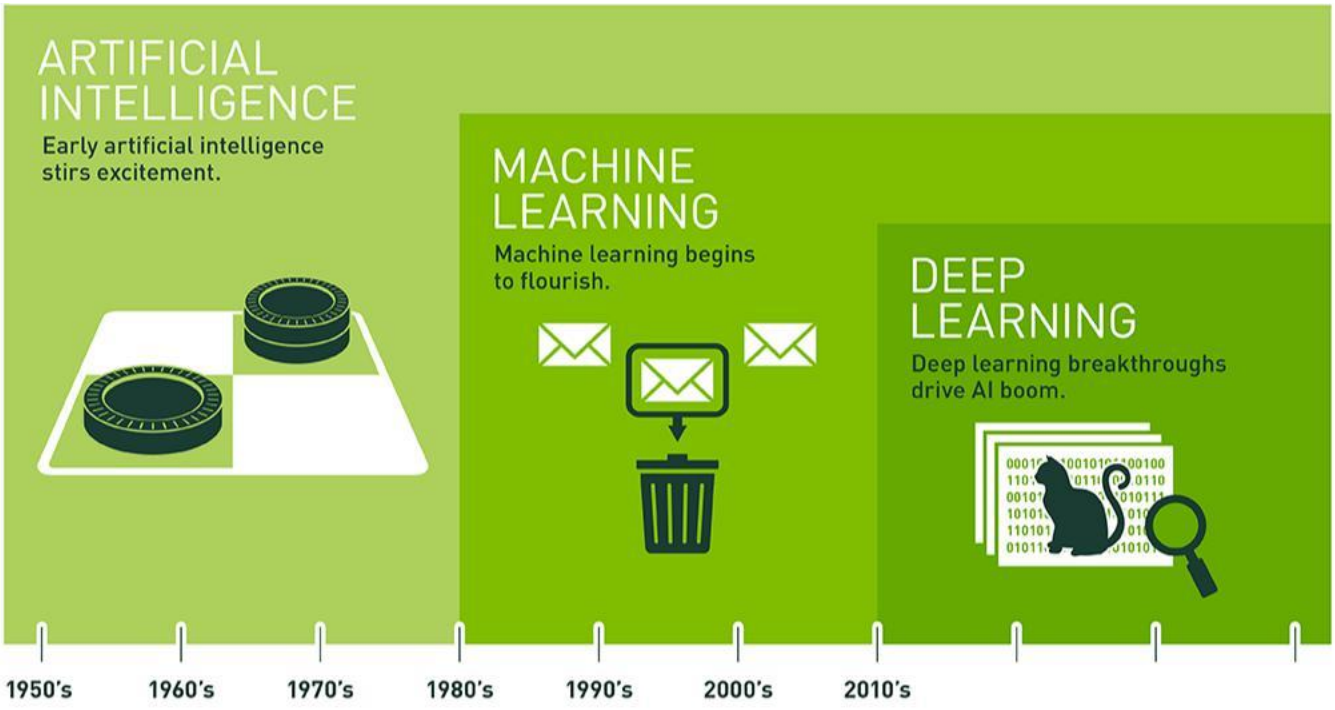
لماذا اداء تعلم العميق جيد و بعض من تاريخه

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" لكن في الآونة الأخيرة فقط بدأنا بالفعل في خدش سطح إمكاناتها الكاملة.

كما وصف Andrej 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!)



Since an early flush of optimism in the 1950s, smaller subsets of artificial intelligence – first machine learning, then deep learning, a subset of machine learning – have created ever larger disruptions.

رسم توضيحي من NVIDIA ، صانع رائد لوحدات معالجة الرسومات (GPUs) التي تم تصميمها في الأصل للألعاب ولكن تبين أنها مناسبة تمامًا لنوع الحوسبة المتوازية التي تتطلبها الشبكات العصبية العميقة

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

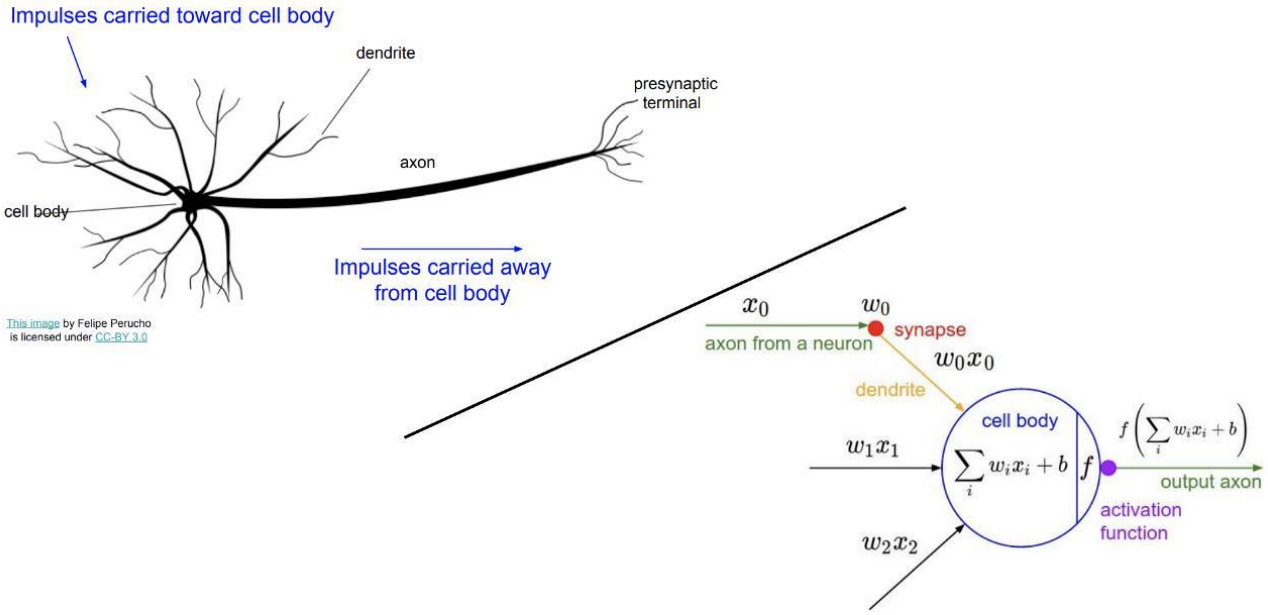
بينما تقرأ هذه الكلمات ، فأنت لا تفحص كل حرف من كل كلمة ، أو كل بكسل يكون كل حرف ، لكي تعرف معنى الكلمة. أنت تجرد

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

> ه ذ ه ال مي زات عا لية الم ست وى هي التي تس مح لك ب ف هم م ا يح دث في هذه ال جم لة دو ن الك ثير من المتا عب (أ و ر ب م ا ت رس ل الكثير م ن الن ص وص في ح الة س كر).

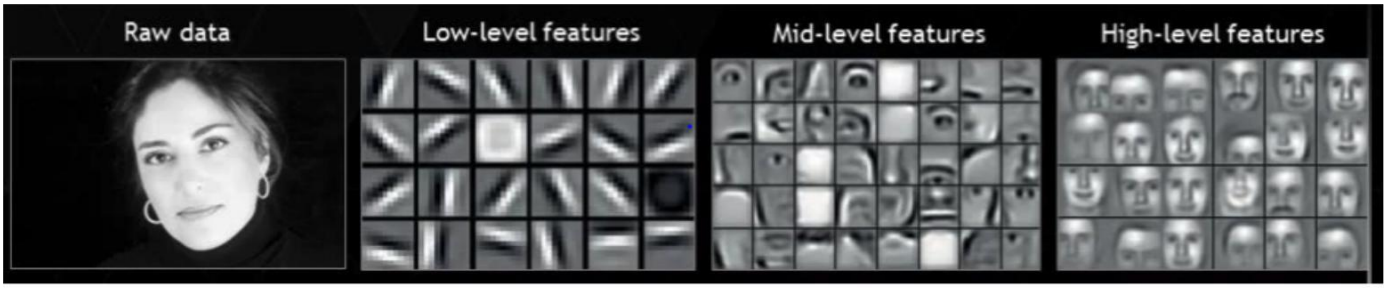
يحدث الشيء نفسه في الرؤية ، ليس فقط في البشر ولكن في النظم البصرية للحيوانات بشكل عام.

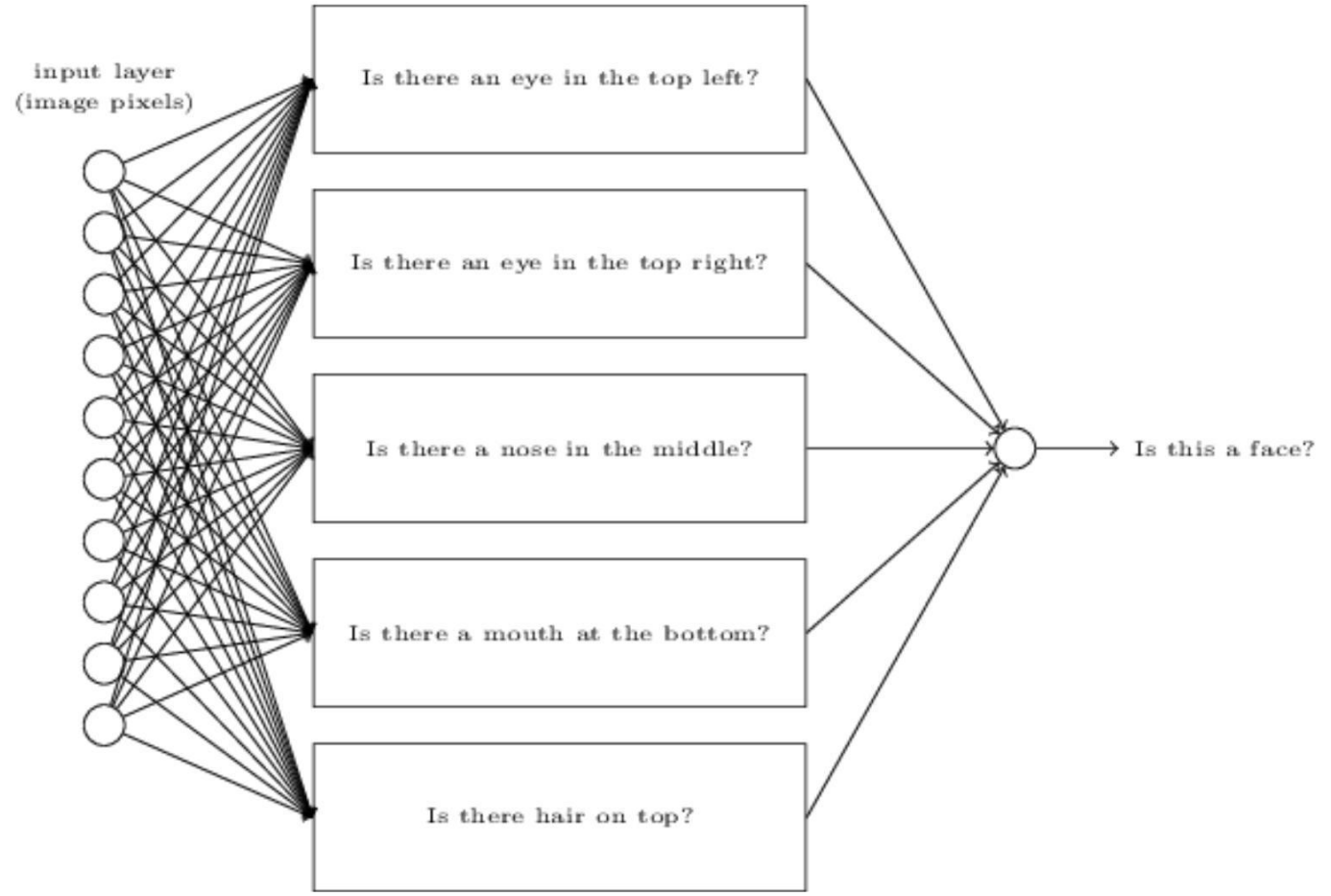
تتكون الأدمغة من الخلايا العصبية التي "تطلق" عن طريق إرسال إشارات كهربائية إلى الخلايا العصبية الأخرى بعد "تنشيطها" بشكل كافٍ. هذه الخلايا العصبية قابلة للطرق من حيث مقدار الإشارة من الخلايا العصبية الأخرى التي ستضيف إلى مستوى تنشيط الخلايا العصبية (بمعنى غامض ، ينتهي الأمر بتدريب الأوزان التي تربط الخلايا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



التوضيحية جنبًا إلى جنب للخلايا العصبية البيولوجية والاصطناعية ، عبر CS231n من جامعة ستانفورد. لا يمكن أخذ هذا التشبيه بشكل حرفي - يمكن للخلايا العصبية البيولوجية أن تفعل أشياء لا تستطيع الخلايا العصبية الاصطناعية القيام بها ، والعكس صحيح - ولكن من المفيد فهم الإلهام البيولوجي. راجع وصف ويكيبيديا للخلايا العصبية البيولوجية مقابل الخلايا العصبية الاصطناعية لمزيد من التفاصيل.

يتم ترتيب شبكاتنا البيولوجية بطريقة هرمية ، بحيث تنتهي بعض الخلايا العصبية باكتشاف ليس ميزات محددة للغاية للعالم من حولنا ، ولكن بالأحرى ميزات أكثر تجريدية ، أي أنماط أو مجموعات من ميزات منخفضة المستوى. على سبيل المثال ، منطقة الوجه المغزلي في النظام البصري للإنسان متخصصة في التعرف على الوجه.





أعلى الصفحة: رسم توضيحي لتعلم الميزات المجردة بشكل متزايد ، عبر NVIDIA. أسفل: رسم تخطيطي لكيفية أخذ الشبكة العصبية الاصطناعية لمدخلات البكسل الخام ، وتطوير "الخلايا العصبية" الوسيطة لاكتشاف الميزات ذات المستوى الأعلى (مثل وجود الأنف) ، والجمع بين مخرجات هذه لإنشاء مخرجات نهائية. رسم توضيحي من الشبكات العصبية والتعلم العميق

(Nielsen, 2017).

تم اكتشاف هذا الهيكل الهرمي الذي عرضته الشبكات العصبية البيولوجية في الخمسينيات من القرن الماضي عندما كان الباحثان David Hubelو 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

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

للاستفادة من دورة التعلم العميق الممتازة في جامعة ستانفورد ، CS231n: الشبكات العصبية التلافيفية والتعرف البصري ، تخيل أننا نريد تدريب شبكة عصبية لتصنيف الصور باستخدام أحد التسميات الصحيحة التالية: "الطائرة" ، "السيارة" ، "الطيور" ، "قط" ، "غزال" ، "كلب" ، "ضفدع" ، "حصان" ، "سفينة" ، "شاحنة".

يمكن أن يتمثل أحد الأساليب في إنشاء "نموذج" ، أو صورة متوسطة ، لكل فئة من فئات الصور باستخدام أمثلة التدريب ، ثم استخدام خوارزمية الجيران الأقرب أثناء

اختبار لقياس مسافة قيم البكسل لكل صورة غير مصنفة ، بشكل إجمالي ، لكل قالب. هذا النهج لا ينطوي على طبقات من التجريد. إنه نموذج خطي يجمع كل الاتجاهات المختلفة لكل نوع من الصور في تمويه متوسط واحد.

على سبيل المثال ، سيستغرق الأمر جميع السيارات - بغض النظر عما إذا كانت تواجه اليسار ، أو اليمين ، أو الوسط ، وبغض النظر عن لونها - ومتوسطها. ثم ينتهي القالب بمظهر غامض وضبابي إلى حد ما.

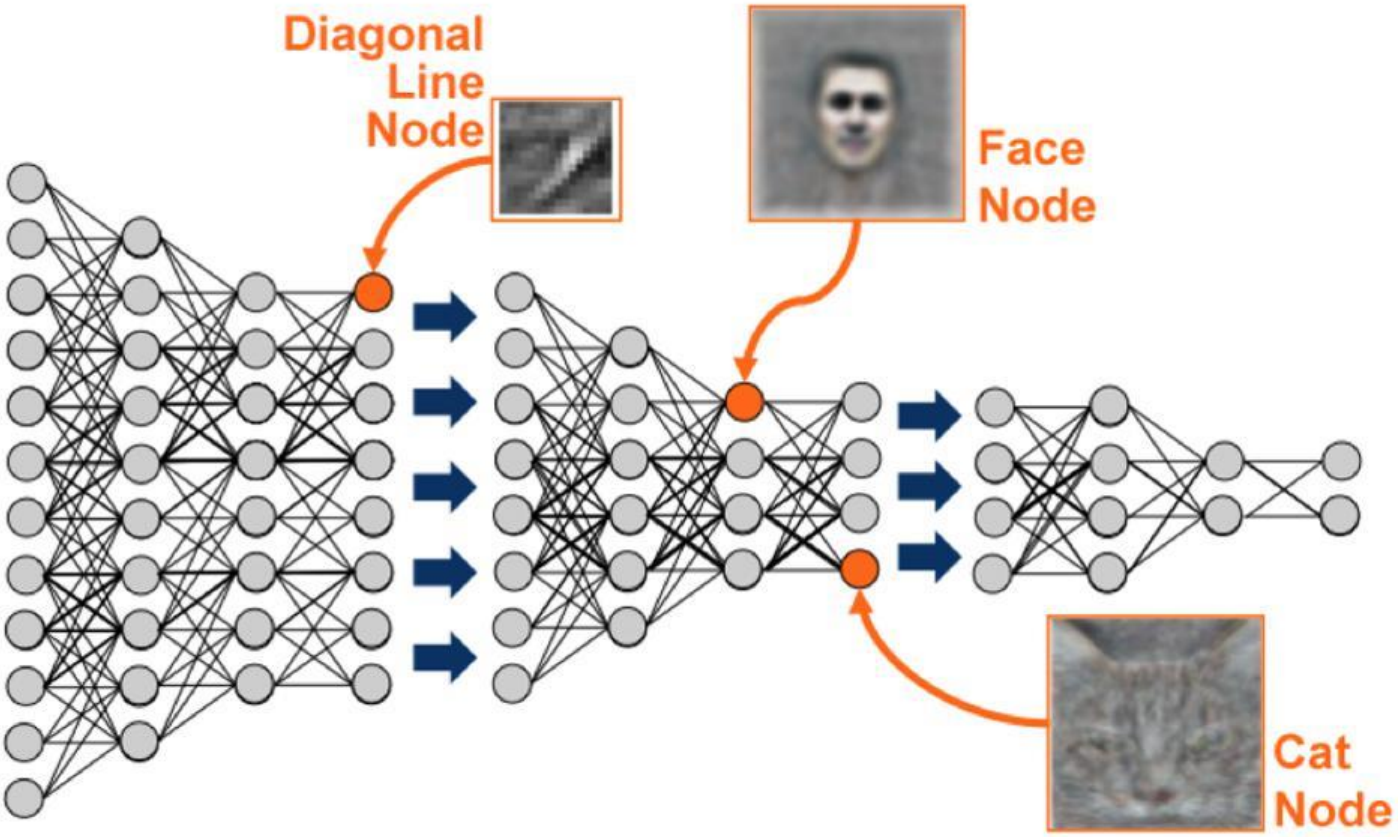
Example drawn from Stanford's CS231n: Convolutional Neural Networks and Visual Recognition, Lecture 2.

لاحظ أن قالب الحصان أعلاه يبدو أنه يحتوي على رأسين. هذا لا يساعدنا حقًا: نريد أن نكون قادرين على اكتشاف الحصان المواجه لليمين أو الحصان الذي يواجه اليسار بشكل منفصل ، وبعد ذلك إذا تم اكتشاف أي من هاتين الميزتين ، فنحن نريد أن نقول إننا ننظر إلى حصان . يتم توفير هذه المرونة من خلال الشبكات العصبية العميقة ، كما سنرى في القسم التالي.

مقاربة الشبكات العصبية العميقة لمشكلة تصنيف الصور باستخدام طبقات التجريد

لتكرار ما شرحناه سابقًا في هذا القسم: ستأخذ طبقة الإدخال سطوع بكسل خام للصورة. ستكون الطبقة الأخيرة عبارة عن متجه ناتج عن احتمالات الفئة (أي احتمال أن تكون الصورة "قطة" أو "سيارة" أو "حصان" ، إلخ.)

ولكن بدلاً من تعلم نموذج خطي بسيط يتعلق بالمدخلات والمخرجات ، سنقوم بدلاً من ذلك ببناء طبقات مخفية وسيطة من الشبكة ستتعلم على نحو متزايد ميزات مجردة ، والتي تمكننا من عدم فقد كل الفروق الدقيقة في البيانات المعقدة.



Source: Analytics Vidhya

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.

يأتي هذا على حساب قابلية تفسير النموذج ، نظرًا لأنه كلما أضفت طبقات مخفية أكثر ، تبدأ الخلايا العصبية في تمثيل المزيد والمزيد من الميزات المجردة وغير المفهومة في النهاية - لدرجة أنك قد تسمع التعلم العميق المشار إليه باسم "تحسين الصندوق الأسود" ، حيث في الأساس ، مجرد تجربة الأشياء بشكل عشوائي إلى حد ما ورؤية ما يخرج ، دون فهم ما يحدث في الداخل حقًا.

. 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.

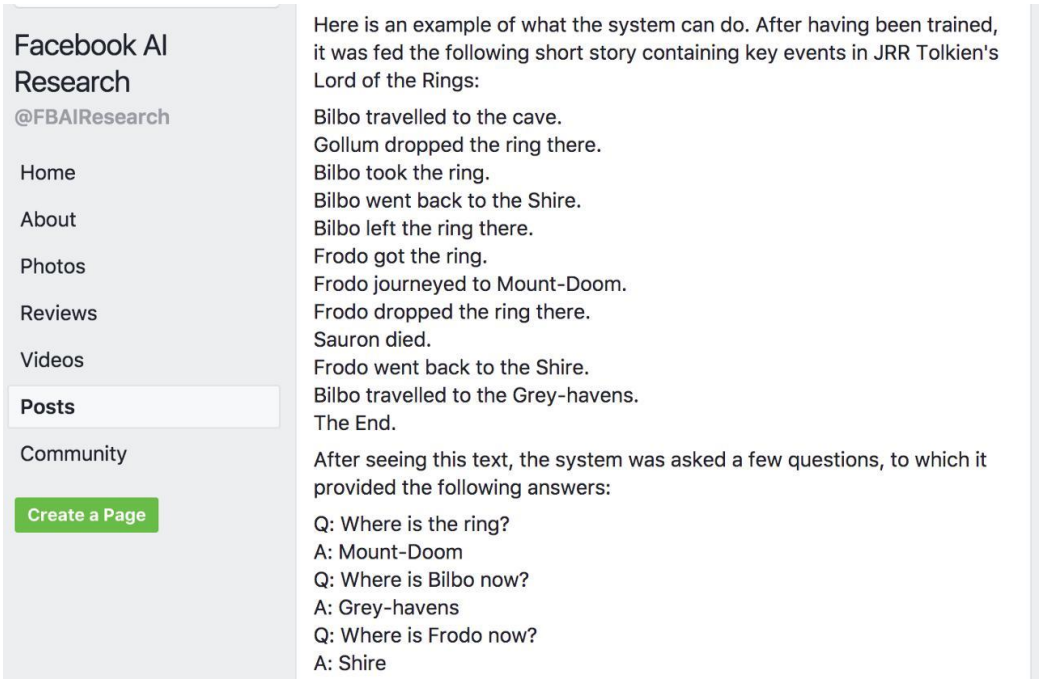
بعض الامتدادات والمفاهيم الإضافية التي تستحق الذكر

- **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 أولاه مجموعة ممتازة من RNNs و LSTMs في سياق مشاكل اللغة.
- •التعلم المقوى العميق. هذا هو أحد أكثر المجالات إثارة في أبحاث التعلم العميق ، في قلب الإنجازات الأخيرة مثل هزيمة OpenAIلاعبي Dota 2 المحترفين و AlphaGo من DeepMind الذي يتفوق على البشر في لعبة .Go سنتعمق أكثر في الجزء 5 ، ولكن الهدف الأساسي هو تطبيق جميع الأساليب في هذا المنشور على مشكلة تعليم الوكيل لزيادة المكافأة إلى أقصى حد. يمكن تطبيق ذلك في أي سياق يمكن تطبيقه - من الألعاب الفعلية مثل Counter Strike أو Pacman ، إلى السيارات ذاتية القيادة ، إلى تداول الأسهم ، إلى (في النهاية) الحياة الواقعية والعالم الحقيقي.

تطبيقات تعلم العميق

التعلم العميق يعيد تشكيل العالم في كل مجال تقريبًا. فيما يلي بعض الأمثلة على الأشياء المدهشة التي يمكن أن يفعلها التعلم العميق...

• قام Facebook بتدريب شبكة عصبية معززة بذاكرة قصيرة المدى للإجابة بذكاء على الأسئلة حول حبكة Lord of the Rings.



. بحث من FAIR (Facebook AI Research) يطبق شبكات عصبية عميقة معززة بذاكرة منفصلة قصيرة المدى للإجابة بذكاء على الأسئلة حول قصة LOTR.هذا هو تعريف الملحمة.

تعتمد السيارات ذاتية القيادة على التعلم العميق للمهام المرئية مثل فهم إشارات الطريق واكتشاف الممرات والتعرف على العوائق

Source: Business Insider

يمكن استخدام التعلم العميق للأشياء الممتعة مثل توليد الفن. يمكن لأداة تسمى النمط العصبي محاكاة أسلوب الفنان بشكل مثير للإعجاب واستخدامها لإعادة مزج صورة أخرى.

The style of Van Gogh's Starry Night applied to a picture of Stanford's campus, via Justin Johnson's neural style implementation: https://github.com/jcjohnson/neural-style

تشمل الأمثلة الأخرى الجديرة بالملاحظة:

- توقع النشاط الحيوي للجزيء لاكتشاف الدواء
- *التعرف على الوجوه والكائن لوضع علامات على الصور والفيديو*
- *تعزيز نتائج بحث Google*
- فهم اللغة الطبيعية وتوليدها ، على سبيل المثال جوجل المترجم
- *The Mars explorer robot Curiosity is* autonomously selecting inspection-worthy soil targets *based on visual examination* …and many, many, more.

الآن اذهب افعلها!

لم ندخل في الكثير من التفاصيل هنا حول كيفية إعداد الشبكات العصبية عمليًا لأنه من الأسهل بكثير فهم التفاصيل من خلال تنفيذها بنفسك. إليك بعض الموارد العملية الرائعة للبدء.

- Play around with the architecture of neural networks to see how different configurations affect network performance with the Google's Neural Network Playground.
- Get up-and-running quickly with this tutorial by Google: TensorFlow and deep learning, without a PhD. 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.

الموارد اخرى

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, Andrew Ng's new deep learning course with a comprehensive syllabus on the subject
- CS231n: Convolutional Neural Networks for Visual Recognition, 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—foundational, more mathematical
- 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

Next up: time to play some games!

Last, but most certainly not least, is Part 5: Reinforcement 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 (رما) البيئة الكلاسيكي لمهام (تم)*
- ***تعلم-كيو ، تعلم السياسة و تعلم المقوى العميق***
- ***وأخيراً ، مشكلة تعلم القيمة***

لنضع فأرً الياً في متاهة

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

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.

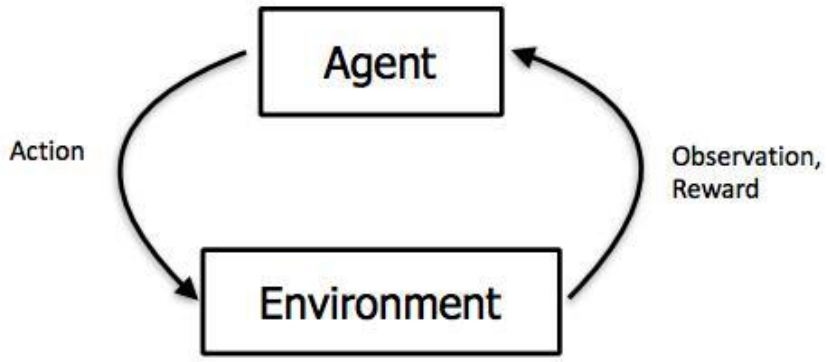
بعد قليل من الاستكشاف ، قد يجد الفأر واحة مصغرة وهي عبارة عن ثلاثة مصادر مائية متجمعة بالقرب من المدخل ، ويقضي كل وقته في استغلال هذا الاكتشاف من خلال الحصول على المكافآت صغيرة لمصادر المياه هذه باستمرار وعدم الذهاب إلى مزيد من المتاهة لمتابعة الجائزة الأكبر.

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.

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 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***
3.

*3. **الانتقال بين الحالات** .على سبيل المثال ، إذا ذهبت يسارًا عند مفترق طرق ، ينتهي بك الأمر في وضع جديد. يمكن أن تكون هذه مجموعة من الاحتمالات التي ترتبط بأكثر من حالة واحدة ممكنة (على سبيل المثال ، عندما تستخدم هجومًا في لعبة بوكيمون ، يمكنك إما أن تفوتك أو تلحق بعض الضرر أو تلحق ضررًا كافيًا لضرب خصمك).*

4. المكافآت المرتبطة بكل انتقال. في مثال فأر الروبوت ، تكون معظم المكافآت صفراً ، لكنها تكون إيجابية إذا وصلت إلى نقطة بها ماء أو جبن وسلبية إذا وصلت إلى نقطة بها صدمة كهربائية.

4. عامل خصم γ بين 0 و 1. يحدد هذا الفرق في الأهمية بين المكافآت الفورية والمكافآت المستقبلية. على سبيل المثال ، إذا كانت γ تساوي 0.9 ، وكانت هناك مكافأة قدرها 5 بعد 3 خطوات ، فإن القيمة الحالية لتلك المكافأة هي $9^3 * 5$.

*4. **بدون ذاكرة** .. بمجرد معرفة الحالة الحالية ، يمكن محو سجل تنقلات الفأر عبر المتاهة لأن حالة ماركوف الحالية تحتوي على جميع المعلومات المفيدة من السجل. بعبارة أخرى ، "المستقبل مستقل عن الماضي بالنظر إلى الحاضر."*

الآن بعد أن عرفنا ما هي (رما) ، يمكننا إضفاء الطابع الرسمي على هدف الفار. نحاول تعظيم مجموع المكافآت على المدى الطويل:

$$\sum_{t=0}^{t=\infty} \gamma^t r(x(t), a(t))$$

دعونا نلقي نظرة على مصطلح هذا المجموع بمصطلح. بادئ ذي بدء ، نقوم بتلخيص جميع الخطوات الزمنية .t دعنا نضبط γ على 1 في الوقت الحالي وننسى الأمر. ص (س ، أ) هي دالة مكافأة. بالنسبة للحالة x والإجراء a (أي الذهاب يسارًا عند مفترق طرق) ، فإنه يمنحك المكافأة المرتبطة باتخاذ هذا الإجراء عند الحالة .x بالعودة إلى معادلتنا ، نحاول تعظيم مجموع المكافآت المستقبلية من خلال اتخاذ أفضل إجراء في كل ولاية.

الآن بعد أن أعددنا مشكلة التعلم المقوى لدينا وقمنا بإضفاء الطابع الرسمي على الهدف ، دعنا نستكشف بعض الحلول الممكنة.

التعلم كيو: تعلم وظيفة قيمة العمل

تعلم-كيو هي تقنية تقوم بتقييم الإجراء الذي يجب اتخاذه بناءً على وظيفة قيمة الفعل التي تحدد قيمة التواجد في حالة معينة واتخاذ إجراء معين في تلك الحالة.

لدينا وظيفة كيو التي تأخذ كإدخال حالة واحدة وإجراء واحد وترجع المكافأة المتوقعة لهذا الإجراء (وجميع الإجراءات اللاحقة) في تلك الحالة. قبل أن نستكشف البيئة ، يعطي كيو نفس القيمة الثابتة (التعسفية). ولكن بعد ذلك ، بينما نستكشف البيئة أكثر ، تعطينا كيو تقديرًا أفضل وأفضل لقيمة الفعل a في حالة .s نقوم بتحديث وظيفتنا كيو مع تقدمنا.

تشرح هذه المعادلة من صفحة ويكيبيديا على تعلم كيو كل شيء بشكل جيد للغاية. يوضح كيف نقوم بتحديث قيمة كيو بناءً على المكافأة التي نحصل عليها من بيئتنا:

$$Q(s_t, a_t) \leftarrow \underbrace{Q(s_t, a_t)}_{\text{old value}} + \underbrace{\alpha}_{\text{learning rate}} \cdot \left(\overbrace{\underbrace{r_t}_{\text{reward}} + \underbrace{\gamma}_{\text{discount factor}} \cdot \underbrace{\max_a Q(s_{t+1}, a)}_{\text{estimate of optimal future value}}}^{\text{learned value}} - \underbrace{Q(s_t, a_t)}_{\text{old value}} \right)$$

دعنا نتجاهل عامل الخصم بتعيينه على 1 مرة أخرى. أولاً ، ضع في اعتبارك أن كيو من المفترض أن تظهر لك المبلغ الكامل للمكافآت من اختيار الإجراء كيو وجميع الإجراءات المثلى بعد ذلك.

الآن دعنا ننتقل إلى المعادلة من اليسار إلى اليمين. عندما نتخذ إجراءً في الحالة st ، نقوم بتحديث قيمة Q (st ، atعن طريق إضافة مصطلح إليها. يحتوي هذا المصطلح على:

- *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*

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).*

الآن بعد أن أصبح لدينا تقدير قيمة لكل زوج من الإجراءات الحكومية ، يمكننا تحديد الإجراء الذي يجب اتخاذه وفقًا لاستراتيجية اختيار الإجراء الخاصة بنا لا نختار بالضرورة الإجراء الذي يؤدي إلى المكافأة الأكثر توقعًا في كل مرة ، على سبيل المثال مع استراتيجية استكشاف (طمع ابسلون) نتخذ إجراءً عشوائيًا في نسبة مئوية من الوقت.

في مثال فأر الروبوت ، يمكننا استخدام تعلم كيو لمعرفة قيمة كل موضع في المتاهة وقيمة الإجراءاتforward} ، backward، left، {rightفي كل موضع. ثم يمكننا استخدام إستراتيجية اختيار الإجراء الخاصة بنا لاختيار ما يفعله الماوس بالفعل في كل خطوة زمنية.

تعلم السياسة: خريطة من القترانات إلى العمل

في نهج تعلم-كيو ، تعلمنا دالة قيمة تقدر قيمة كل زوج من الإجراءات بين الدول.

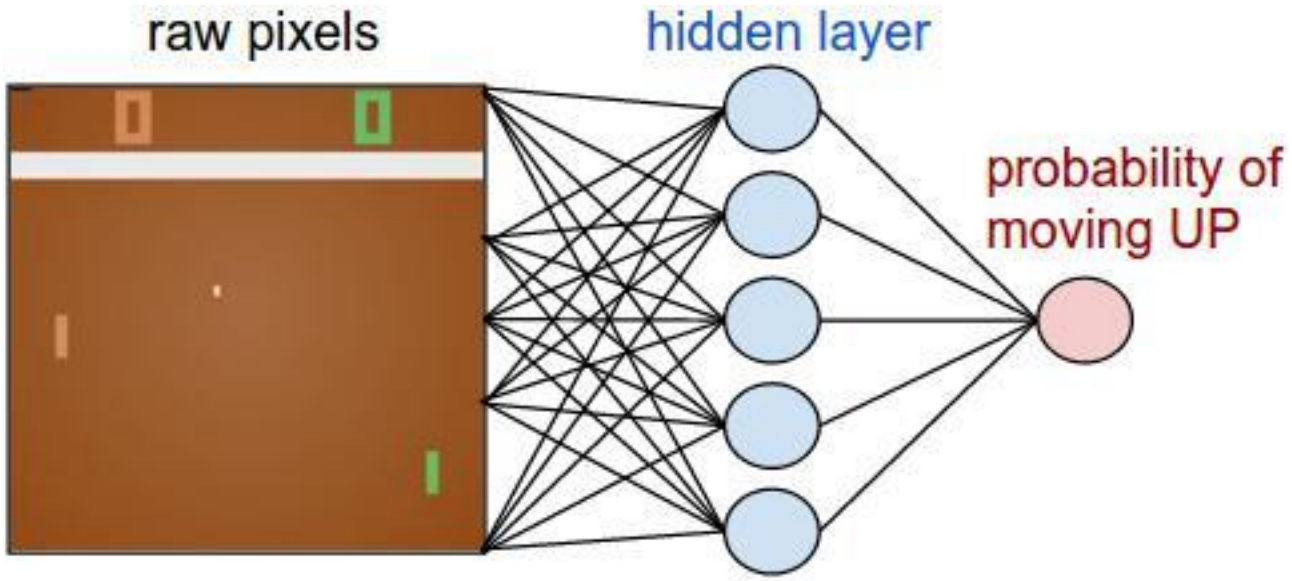
يعد تعلم السياسة بديلاً أكثر وضوحًا حيث نتعلم وظيفة السياسة ، ، وهي خريطة مباشرة من كل ولاية إلى أفضل إجراء مطابق في تلك الحالة. فكر في الأمر كسياسة سلوكية: "عندما ألاحظ الدول ، فإن أفضل شيء أفعله هو اتخاذ إجراء أ". على سبيل المثال ، قد تتضمن سياسة السيارة ذاتية القيادة بشكل فعال شيئًا مثل: "إذا رأيت ضوءًا أصفر اللون وكنت على بعد أكثر من 100 قدم من التقاطع ، يجب أن أكبح. وإلا ، فاستمر في المضي قدمًا."

$$a = \pi(s)$$

السياسة هي عبارة عن تحول من حالة إلى فعل

لذلك نحن نتعلم وظيفة من شأنها زيادة المكافأة المتوقعة. ما الذي نعرف أنه جيد حقًا في تعلم الوظائف المعقدة؟ شبكات عصبية عميقة!

يوفر Andrej Karpathy's Pong from Pixels تجولًا ممتازًا حول استخدام التعلم المعزز العميق لتعلم سياسة لعبة Atari Pong التي تأخذ وحدات البكسل الأولية من اللعبة كمدخل (حالة) وتخرج احتمالية تحريك المضرب لأعلى أو لأسفل (حركة).



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/

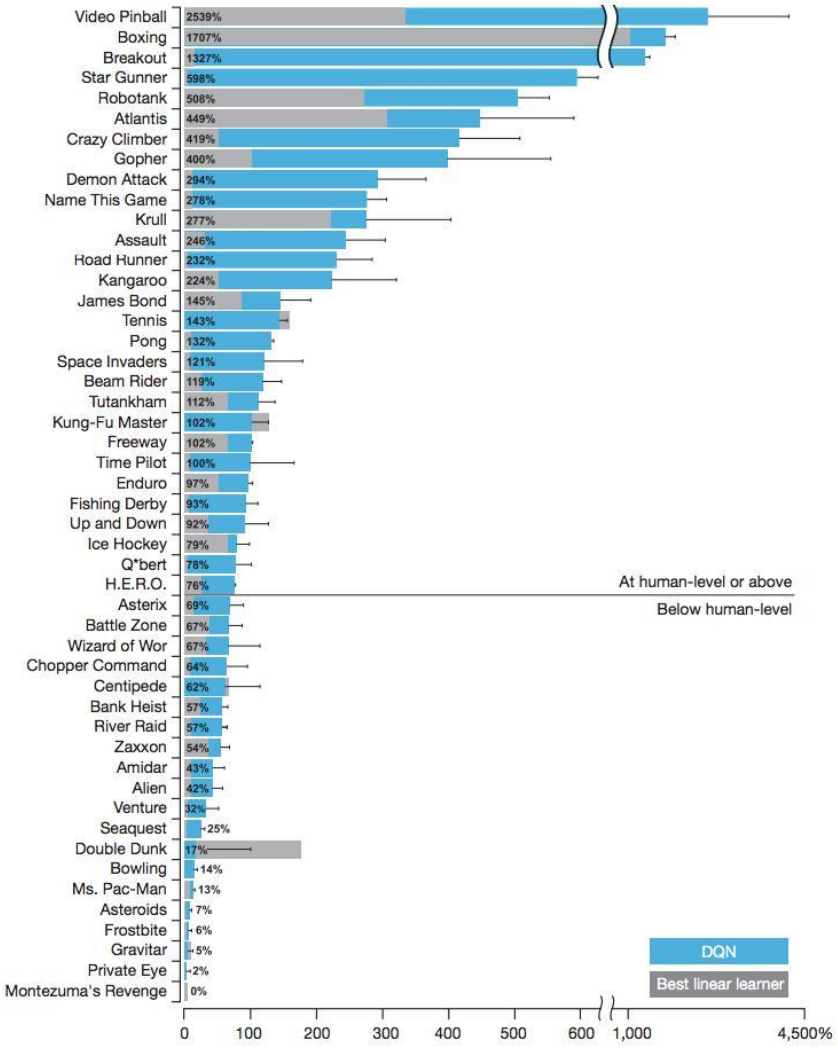
إذا كنت تريد أن تتسخ يديك باستخدام RL العميق ، فاعمل من خلال مشاركة .Andrej ستنفذ شبكة سياسة مكونة من طبقتين في 130 سطرًا من التعليمات البرمجية ، وستتعلم أيضًا كيفية التوصيل بـ OpenAI's Gym ، والذي يسمح لك بالاستعداد والتشغيل بسرعة باستخدام خوارزمية التعلم التعزيزية الأولى الخاصة بك ، واختبارها على مجموعة متنوعة من الألعاب ، ونرى كيف يقارن أدائها بالإرسالات الأخرى.

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:

لقد أثبتنا أن وكيل تعلم –كيو العميق ، الذي يتلقى فقط وحدات البكسل ودرجة اللعبة كمدخلات ، كان قادرًا على تجاوز أداء جميع الخوارزميات السابقة وتحقيق مستوى مماثل لمختبر ألعاب بشرية محترف عبر مجموعة من 49 لعبة ، باستخدام نفس الخوارزمية وبنية الشبكة والمعلمات الفائقة. يعمل هذا العمل على سد الفجوة بين المدخلات والإجراءات الحسية عالية الأبعاد ، مما ينتج عنه أول عامل اصطناعي قادر على تعلم التفوق في مجموعة متنوعة من المهام الصعبة. (سيلفر وآخرون ، 2015)

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

لمساعدتك في بناء حدس لكيفية تحقيق التقدم في أبحاث RL ، إليك بعض الأمثلة على التحسينات على المحاولات في مقاربات اقتران-كيو غير الخطية التي يمكنها تحسين الأداء والاستقرار:

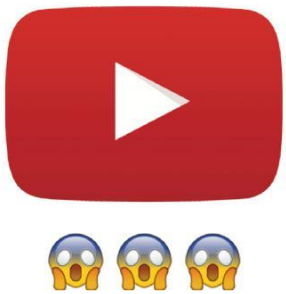
- **Experience replay**, *which learns by randomizing over a longer sequence of previous observations and corresponding reward to avoid overfitting to recent experiences. This idea is* inspired by biological brains*: rats traversing mazes, for example, "replay" patterns of neural activity during sleep 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 *to know things exist even if they leave the baby's visual field. RNNs are "recurrent", i.e.* they allow information to persist *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 with Tensorflow 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 ، A3C 2016). 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.

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 *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*)."*
- *Next, we'd highly recommend Arthur Juliani's* Simple Reinforcement Learning with Tensorflow *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*—a fantastic book, very readable*
- *John Schulman's* CS 294: Deep Reinforcement Learning *(Berkeley)*
- *David Silver's* Reinforcement Learning *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.

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.

الآساسيات

البرمجة

Syntax and basic concepts: Google's Python Class, Learn Python the Hard Way.
Practice: Coderbyte, Codewars, HackerRank.

الجبر الخطي

Deep Learning Book, Chapter 2: Linear Algebra. A quick review of the linear algebra concepts relevant to machine learning.
A First Course in Linear Model Theory 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, 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, by Larry Wasserman. Introductory text on statistics.

تفضال و تكامل

Khan Academy: Differential Calculus. Or, any introductory calculus course or textbook.
Stanford CS231n: Derivatives, Backpropagation, and Vectorization, prepared by Justin Johnson.

تعلم الآلة

دورات

Andrew Ng's Machine Learning course on Coursera (or, for more rigor, Stanford CS229). Machine learning bootcamps: Galvanize (full-time, 3 months, $$$$), Thinkful (flexible schedule, 6 months, $$).

كتب

An Introduction to Statistical Learning by Gareth James et al. Excellent reference for essential machine learning concepts, available free online.

تعلم العميق

الدورات

Deeplearning.ai, Andrew Ng's introductory deep learning course.
CS231n: Convolutional Neural Networks for Visual Recognition, Stanford's deep learning course. Helpful for building foundations, with engaging lectures and illustrative problem sets.

المشاريع

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. Classify handwritten digits with >99% accuracy in 3 hours with this tutorial by Google.
Try your hand at a Kaggle competition. Implement a deep learning paper that you found interesting, using other versions on GitHub as reference material.

المطبوعات

Deep Learning Book, a.k.a. the Bible of Deep Learning, authored by Ian Goodfellow, Yoshua Bengio, and Aaron Courville.
Neural Networks and Deep Learning, a clear and accessible online deep learning text by Michael Nielsen. Ends with commentary on reaching human-level intelligence.
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 at Berkeley.
David Silver's Reinforcement Learning course at University College London.
Deep RL Bootcamp, 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. Implement a Pong-playing agent from scratch in 130 lines of code.

Arthur Juliani's Simple Reinforcement Learning with Tensorflow series. Implement Q-learning, policy-learning, actor-critic methods, and strategies for exploration using TensorFlow.
See OpenAI's requests for research for more project ideas.

المطبوات

Richard Sutton's book, Reinforcement Learning: An Introduction.

الذكاء الاصطناعي

Artificial Intelligence: A Modern Approach by Stuart Russell and Peter Norvig.
Sebastian Thrun's Udacity course, Intro to Artificial Intelligence.
Fellowships: Insight AI Fellows Program, Google Brain Residency Program

سلامة الذكاء الاصطناعي

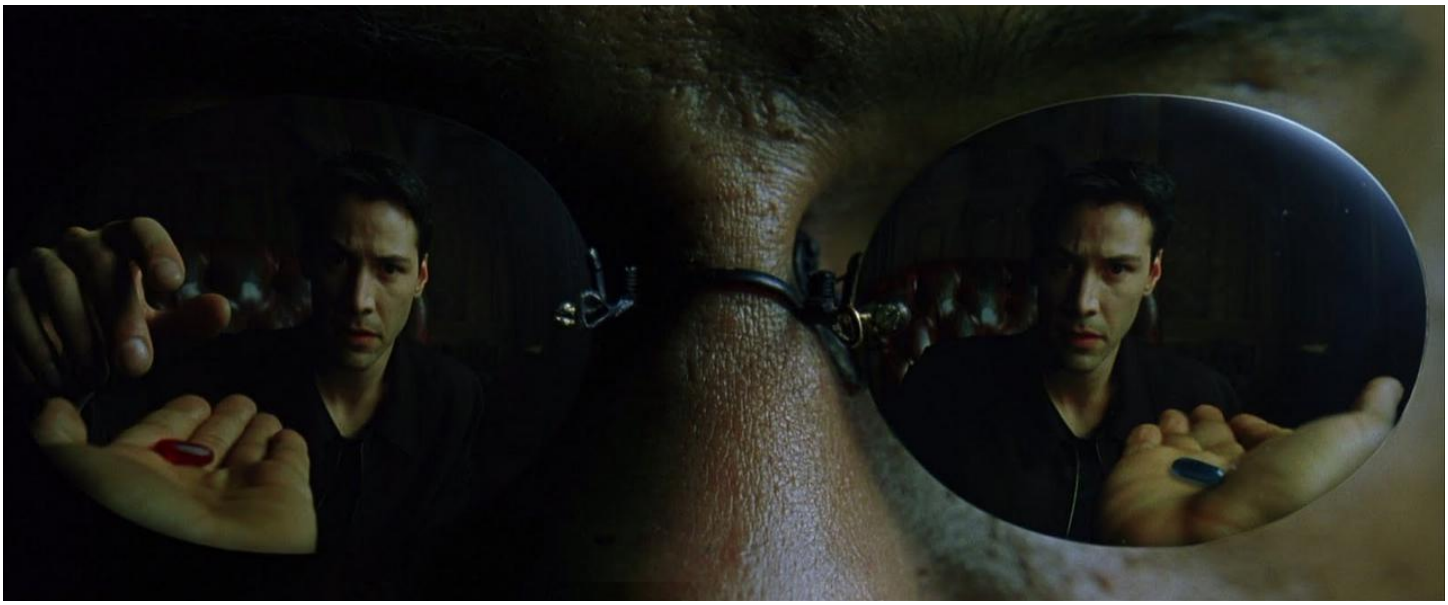
For the short version, read: (1) Johannes Heidecke's Risks of Artificial Intelligence, (2) OpenAI and Google Brain's collaboration on Concrete Problems in AI Safety, and (3) Wait But Why's article on the AI Revolution.
For the longer version, see Nick Bostrom's Superintelligence.
Check out the research published by the Machine Intelligence Research Institute (MIRI) and Future of Humanity Institute (FHI) on AI safety. Keep up-to-date with /r/ControlProblem on Reddit.

مقالات

Import AI, weekly AI newsletter covering the latest developments in the industry. Prepared by Jack Clark of OpenAI.
Machine Learnings, prepared by Sam DeBrule. Frequent guest appearances from experts in the field.
Nathan.ai, covering recent news and commenting on AI/ML from a venture capital perspective.

نصائح

“What is the best way to learn machine learning without taking any online courses? — answered by Eric Jang, Google Brain
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, a more code-based introduction to AI
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