مقدمة في التعلم الآليِّ المفسر

...

الطالب: سَرِيٌّ بنُ محمد تيسير السباعي

هندسة برمجيات في جامعة الأمير سلطان

تَقْدِمةٌ في التَّسْمِيَة

- لم أجد من ترجم هذا المصطلح إلا معجمَ سدايا إذ قال "ذكاء صناعي قابل التفسير" ولكني أحببت عرض بضاعتي في نقاش هذا المصطلح فأقول ان المقصد من التفسير تبيين سبب قرار النموذج فلعلنا ننسبه للتعلم بدل الذكاء فتقول "التعلم الآليّ المفسَّر"

وهذه التسمية والله أعلم مناسبة لنصعد منها الى مرتبة أعلى وهي التعليم المعلل

[وقد كان هذا الاسم هو الراجح عندي لهذا العلم ولكن مخالفة المشهور ضارة ولو نفع بعضها فبقيت على الأصل وغيرته]

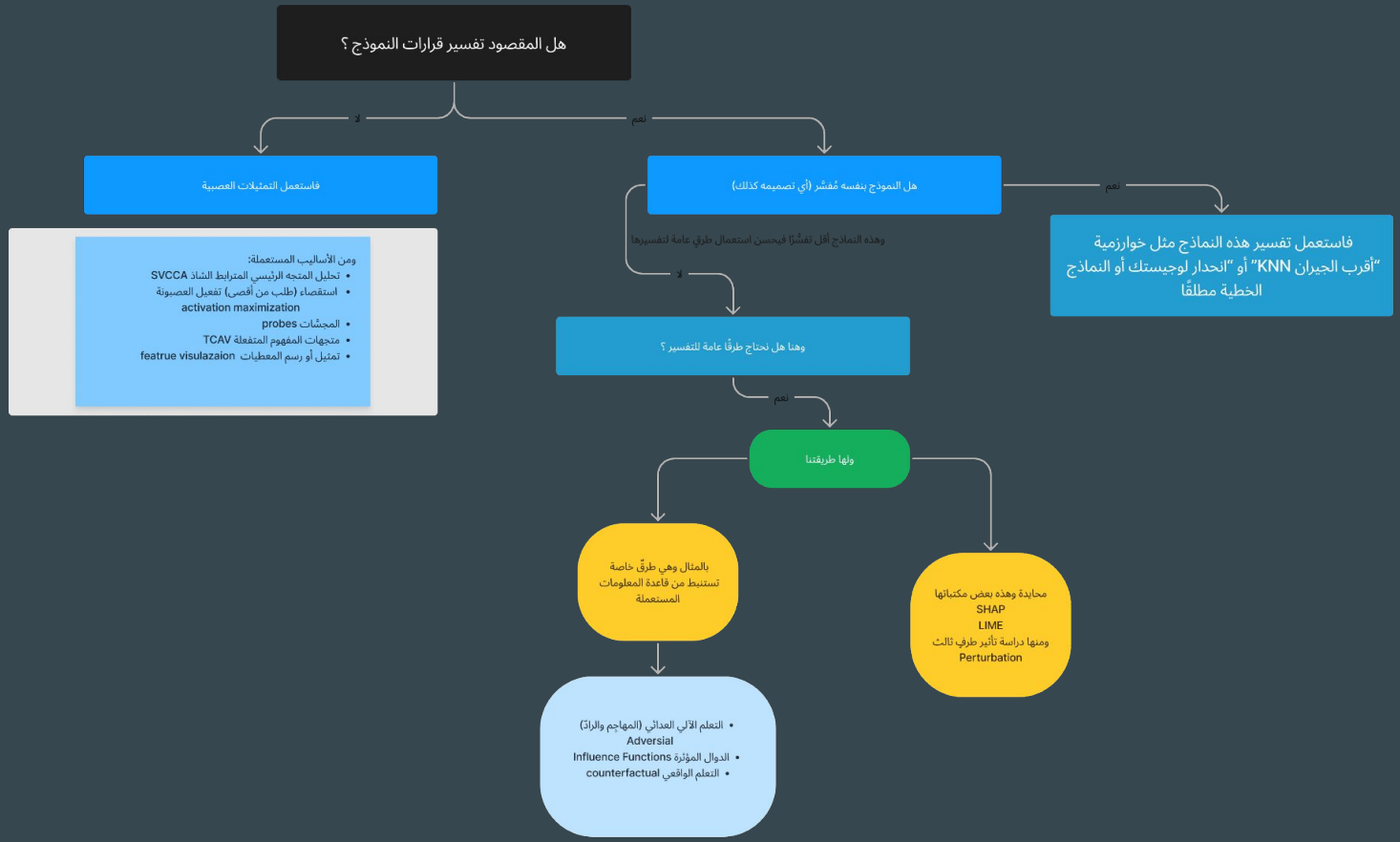
تعريف المجال وسببه

- إن انتشار استعمال الذكاء الصناعي عامَّةً والتعلم الآلي والعميق خاصة في حَلِّ كثيرٍ من المشاكل كان سابقا لتفسير أسبابه ومع توسعه ودخوله في مجالات دقيقة يُتَحسَّسُ فيها عُشيرٌ ومعشار كالطب والمعالِجَات أصبح لزاما وجود طرقٍ تفسّر العمل لكل صغيرة فيما يَتعلم من النماذج وخاصة العميقة منها.
- ولهذا بزغ نجم علم التفسير الآلي الموصل للتعليل الداخلي للأوزان المتعلَّمَة مبتدِاً بالمعادلات الرياضية و خلوصا إلى التطبيق البرمجي ولكلٍّ مثالٌ آتٍ ان شاء الله

أنواعه وطرق الاختيار

- وهذا ملخص استحسنت إيراده شامل بعموم في مجالات الفنّ مذكور في الشريحة التالية

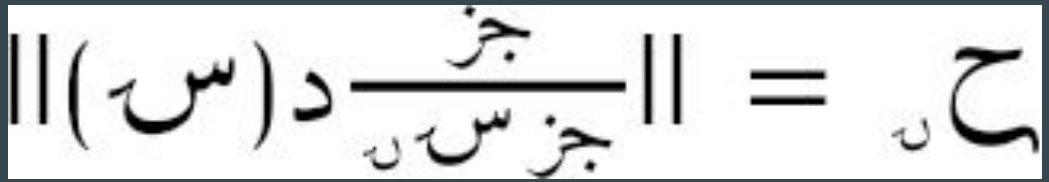
Insert the question or decision you plan on analyzing here



أمثلة رياضية

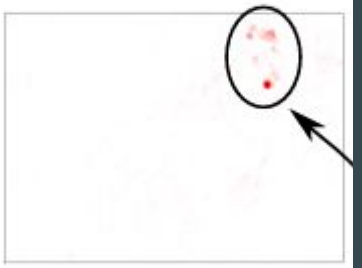
- ولا يشك شاك أن الرياضيات هي أساس هذا المجال و أصله وقاعدته ولهذا لا بد لسالكه أن يتضلع منها تضلعه من الماء في يومه فليس الماشي كالعارج ولا بد من الجهد في صعود المعارج وسأذكر إن شاء الله طريقة واحدة من عديدات وهي حاسبة الحساسية

حيث تظهر حساسية تأثير تغيير شيء من المدخلات



$$R_i = \left\| \frac{\partial}{\partial x_i} f(\mathbf{x}) \right\|.$$

المدخل س

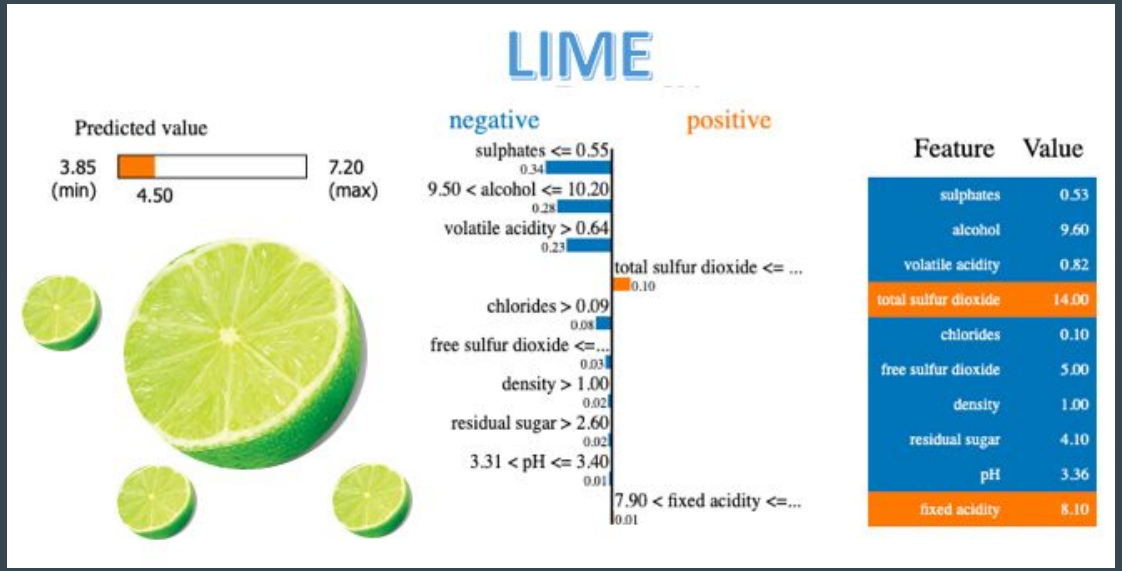


$$ح_{ن} = \left\| \frac{جز}{جز_{س_{ن}}} د(س_{ن}) \right\|$$

والتنبيه هنا أن هذا باعتبار أن كل عنصورة (نقطة) مدخل لوحده وهذا نموذج قديم ولكن عليه المثال

أمثلة مُبَرْمَجَةٌ

- وسنستخدم ان شاء الله مكتبة LIME وهناك أيضا مكتبة SHAP ولعلها للقاء آخر ان شاء الله



مثال سريع

- وهذا مثال استعملنا فيه المكتبة لنرى أي المواضع التي رُكِّزَ عليها في حكم النموذج

```
from lime import lime_image
import time
```

```
explainer = lime_image.LimeImageExplainer()
```

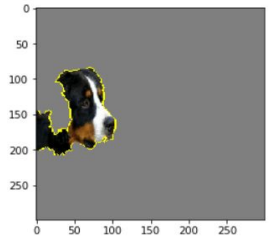
Now let's see the explanation for the top class (Bernese mountain dog) ¶

We can see the top 5 superpixels that are most positive towards the class with the rest of the image hidden

```
from skimage.segmentation import mark_boundaries
```

```
temp, mask = explanation.get_image_and_mask(240, positive_only=True, num_features=5, hide_rest=True)
plt.imshow(mark_boundaries(temp / 2 + 0.5, mask))
```

```
<matplotlib.image.AxesImage at 0x116ea4cd0>
```



مثال سريع

- وهذا مثال استعملنا فيه المكتبة لنرى أي المواضع التي رُكِّزَ عليها في حكم النموذج

```
from lime import lime_image
import time
```

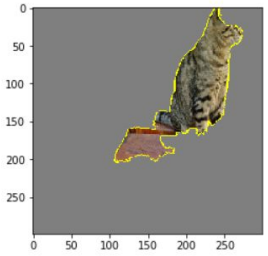
```
explainer = lime_image.LimeImageExplainer()
```

Let's see the explanation for Egyptian cat

Most positive towards egyptian cat:

```
temp, mask = explanation.get_image_and_mask(286, positive_only=True, num_features=5, hide_rest=True)
plt.imshow(mark_boundaries(temp / 2 + 0.5, mask))
```

```
<matplotlib.image.AxesImage at 0x1209e1450>
```



الخاتمة

- وبعد فهذه أهم العناصر مما يحسن التنبيه عليه في الختام
  - أن هذا العلم مازال جنينا فيما كُتبَ فيه بلساننا (ولعل هذه الورقة فاتحة خير) ويافعا بلغة الانجليز
  - أن التفسير مهم وخاصة في المجالات الحساسة
  - أن اشتقاق المصطلحات في العربية ينبغي التحقيق والتحرير فيه
  - أن كل قسم مما ذكر يمكن شرحه في ساعات ودروس
- وأهم مشكلة قابلتني أثناء البحث أني لم أجد أي شيء مكتوب عن المجال فاضطررت لترجمة كل شيء

خلاصة

بسم الله والحمد لله أما بعد: فهذا بحث عن الذكاء والتعلم المُفَسَّرِ أو القابل للتفسير اذ تأتي أهميته تابعة للإنتشار المتفاقم لتطبيقات تعلم الآلة والتعلم العميق في عموم المجالات الصناعية والبحثية في السنوات الماضية فكان لا بد من أساليب واضحة وسبل رصينة متينة تبين أسباب عمل هذه التقنيات الحديثة فما من عادة الإنسان1 قبول مالا يدري سببه أو استعماله. فكان الغرض من هذه الورقات تقديم تمهيد في هذا الباب باللغة العربية يكون مقسما على التوالي من الأقسام فالأول منها في تعريف هذا المصطلح وتبيينه وذكر مكانه من مجال تعلم الآلة والتعلم العميق مع تبيين بعض غوامضه واشكالاته والثاني فذكر التقنيات المستعملة فيه وطريقة اختيارها حسب كل فرع من فروع تعلم الآلة مرتَّبًا من أسهلها غير المعقد وصولا الى أشدّها تعقيدا وعمقا كالخلايا العصبية عديدة الطبقات ومن ثَمّ في الباب الثالث نذكر تطبيقا عمليا لعدد من التقنيات البرمجية لنماذج حقيقية تم تدريبها ومحاولة تفسيرها باستعمال طرق كثيرة المشتقات المتكاملة وغيرها مراعيا فيه تعقيد النموذج اذ نبدأ فيه من الواضح وصولا لشديد التعقيد وسنذكر من كل استعمال ما يُفسِّره ويظهره من النتائج ان شاء الله. ولعلي أشير الى منهجي المتواضع في كتابة هذه الوُريقات اذ هو مبنيٌّ على تتبع المصادر الأجنبية في المجال وتلخيص نتائجها لانعدام أي مصدر عربي عن هذا العنوان المطروق [عندهم] ولعل هذه أول ورقة فيه بلسان العرب. وأما نتيجة البحث فهي تتلخص في استيضاح أهمية هذا الباب من علم تعلم الآلة والتعلم العميق اذ لا بد للمهندس والعالِم فيه أن يكون قادرا على تفسير نتائج نموذجه المبْنِيّ وتعليل مخرجاته وإلا لصار تابعا أعمى لما صنعت يداه. وكذا وجب التنبيه الى شح الشارحين لهذا المجال بالعربية ولعل هذه الورقة مفتاح خير لهذا الباب إن شاء الله تعالى.

المصادر

العربية:

- https://sdaia.gov.sa/files/Dictionary.pdf
- معجم الرياضيات (دمشق) https://shortest.link/8BdB
- خط سين للتنضيد العربي https://khatt.org/

المصادر

الإنجليزية:

- https://github.com/marcotcr/lime
- https://iphome.hhi.de/samek/pdf/SamITU18b.pdf
- https://ex.pegg.io/
-