المحاضرة 10

الدرس العاشر: استخراج النصوص وتحليل بياناتها - استخراج العلاقات والتحليل المتقدم

في هذا الجزء الثاني، سننتقل من مجرد التظليل (NER) إلى الفهم الهيكلي للنص. سنحول الجمل الوصفية الطويلة إلى خرائط معرفية توضح الروابط بين العناصر التربوية المختلفة.[1]

1. استخراج العلاقات (Relation Extraction - RE)

استخراج العلاقات هو العملية التي تلي التعرف على الكيانات. هدفها تحديد الرابط المنطقي بين كيانين داخل الجملة.

في سياقنا التربوي: إذا وجدنا اسم أستاذ خالد (شخص) واسم مادة الرياضيات (موضوع)، فإن استخراج العلاقة يخبرنا بأن الأستاذ خالد يُدرّس الرياضيات.[3]

المكونات الثلاثية للعلاقة (Triple): تتكون أي علاقة نستخرجها من:

المبتدأ (Subject): الكيان الأول.

العلاقة (Relation): الفعل أو الرابط.

الخبر/المفعول (Object): الكيان الثاني.[2]

■ ***مثال:*** أظهرت ليلى مهارة متميزة في الرسم.

- بايثون تستخرج: (ليلى ← تمتلك مهارة ← الرسم).

2. تحليل التبعية (Dependency Parsing)

لفهم العلاقات بدقة، تستخدم بايثون تقنية تسمى تحليل التبعية. هي عملية ترسم شجرة توضح كيف ترتبط الكلمات ببعضها نحوياً.[2]

تساعدنا هذه الشجرة في معرفة أي صفة تعود على أي اسم.

■ **التطبيق:** في التقارير النفسية المعقدة، تساعدنا هذه التقنية في فض الاشتباك بين الجمل؛ فنعرف يقيناً أن صفة الخمول كانت توصف بها البيئة الصفية وليس الطالب نفسه.[2]

3. المستويات المتقدمة لتحليل المشاعر (Sentiment Analysis)

في الدرس السابق، عرفنا أن التحليل قد يكون إيجابياً أو سلبياً. اليوم، سنتعرف على مستويات أكثر عمقاً:

تحليل المشاعر على مستوى الجملة (Sentence-level): بدلاً من تقييم النص كاملاً، نقيم كل جملة على حدة. قد يكون تقرير المعلم إيجابياً في البداية (أحمد طالب مجتهد) وسلبياً في النهاية (لكنه يتنمر على زملائه).

تحليل المشاعر القائم على الجوانب (Aspect-based): وهو الأهم لنا؛ حيث نحدد مشاعر الطالب تجاه جوانب محددة. (مثلاً: مشاعر إيجابية تجاه المادة العلمية لكن مشاعر سلبية تجاه طريقة التقييم).

اكتشاف العواطف العميقة: استخدام نماذج برمجية لاكتشاف (الفرح، الحزن، الغضب، القلق، الإحباط) بدلاً من مجرد (+) أو (-).[5]

4. التصوير البصري للكيانات والعلاقات (Visualization)

كأخصائيين، يسهل علينا فهم الرسوم التوضيحية أكثر من الأكواد الجافة. توفر مكتبة **spaCy** أداة تسمى displacy تقوم برسم العلاقات والكيانات داخل النص بأسلوب بصري جذاب.

تقوم الأداة بتلوين الأسماء (الأشخاص بلون، والمؤسسات بلون آخر).

ترسم أسهماً توضح الروابط النحوية بين الكلمات.[4]

5. تحليل الشبكات الاجتماعية (Network Analysis) داخل الفصل

باستخدام استخراج العلاقات، يمكننا بناء شبكة اجتماعية للفصل الدراسي.

إذا حللنا نصوص من صديقك المفضل؟ التي يكتبها الطلاب، يمكن لبايثون رسم شبكة توضح:

الطلاب المركزيين (الأكثر شعبية).

الطلاب المنعزلين (الذين لم يذكرهم أحد)، وهم الفئة التي نحتاج للتركيز عليها تربوياً.

6. أدوات عملية (Libraries) للتحليل المتقدم

للقيام بهذه المهام المعقدة، نعتمد على:

spaCy: ملكة استخراج الكيانات والعلاقات والتحليل النحوي.[7]

TextBlob: مكتبة بسيطة وسريعة لتحليل المشاعر والترجمة.

VADER: مكتبة متخصصة في تحليل المشاعر في النصوص القصيرة (مثل منشورات الطلاب على المنصات التعليمية).[8]

7. الربط بالواقع: معالجة سجلات السلوك

تخيل أنك تعمل في مدرسة بها 1000 طالب، ولديك سجلات سلوك يومية يكتبها 50 معلماً.

يدوياً: مستحيل قراءتها يومياً واكتشاف الأنماط.

بايثون (2 IE Part):

تستخرج أسماء الطلاب المتكررين في المخالفات.

تستخرج نوع العلاقة (شجار، تأخر، غياب).[3]

تحلل نبرة المعلمين (هل هناك معلم نبرته سلبية دائماً تجاه الطلاب؟).

تعطيك إنذاراً مبكراً قبل وقوع مشكلة أكبر.[1]

الخلاصة

يا أبنائي، إن استخراج النصوص في جزئه الثاني ينقلكم من مرحلة القراءة إلى مرحلة الاستبصار (Insight). أنتم الآن قادرون على رؤية الروابط الخفية بين الكلمات وتحويل المشاعر الذاتية إلى أرقام ومخططات موضوعية.

■ **تذكر دائماً:** الحاسوب يحلل الكلمات، لكنك أنت من يفسر المعنى النفسي والتربوي العميق خلف هذه النتائج.

في الدرس القادم، سننتقل إلى موضوع تقني هام جداً يسمى **توحيد النصوص (Text Unification)**؛ حيث سنتعلم كيف نجعل الحاسوب يدرك أن أحمد، واحمد، وأحمد (بالفراغ) هم شخص واحد، لضمان دقة إحصاءاتنا.[1]

مناقشة:

لو أعطيتك تقريراً يقول: يشعر الطالب بالخوف من مادة الرياضيات بسبب صعوبة الاختبارات. حاول استخراج (المبتدأ - العلاقة - المفعول به) يدوياً قبل أن نطبقها برمجياً.