المحاضرة 9

مهارات الحاسوب ومعالجة النصوص

المرتكزات التقنية والمعرفية لطلبة علم النفس التربوي

تمثل مهارات الحاسوب ومعالجة النصوص حجر الزاوية في التكوين المعاصر لطلبة علم النفس التربوي، حيث لم يعد الحاسوب مجرد أداة مكتبية ثانوية، بل صار امتداداً للعمليات المعرفية والبحثية التي يمارسها الأخصائي في الميدان التعليمي. إن الانتقال من العصر التقليدي إلى العصر الرقمي فرض تحولاً في السوسيوسكوب أو المنظار الاجتماعي الذي نراقب من خلاله السلوك البشري، مما أدى إلى ظهور ما يعرف بالعلوم الاجتماعية الحاسوبية التي تستفيد من القوى الحسابية الهائلة لفهم التفاعلات الإنسانية المعقدة.[1]

(محتوى الدروس من 1 إلى 8 كما ورد سابقاً...)

الدرس التاسع: استخراج النصوص وتحليل بياناتها في بايثون (الجزء الأول)

في هذا الدرس، سنبدأ في تعلم تقنيات **استخراج المعلومات ((Information Extraction - IE**. إذا كانت الدروس السابقة قد علمتنا كيف نخزن البيانات، فإن هذا الدرس سيعلمنا كيف نستخلص المعاني والكيانات من داخل النصوص غير المنظمة، مثل المقابلات الإكلينيكية والتقارير المدرسية.

1. ما هو استخراج المعلومات (IE)؟

استخراج المعلومات هو عملية آلية تهدف إلى تحويل النصوص غير المهيكلة (مثل فقرة مكتوبة) إلى بيانات مهيكلة (مثل جدول يضم الأسماء والتواريخ والتشخيصات). بالنسبة للباحث التربوي، هذا يعني تقليل الوقت الضخم الذي يُقضى في قراءة الملفات يدوياً للبحث عن أنماط معينة.

2. التعبيرات النمطية (Regular Expressions - Regex): المرشح الذكي

تعد التعبيرات النمطية (Regex) واحدة من أقوى الأدوات في بايثون للبحث عن أنماط نصية محددة. هي ليست مجرد بحث عن كلمة، بل بحث عن شكل المعلومة.

- **لماذا نستخدمها؟** للبحث عن أرقام هواتف، تواريخ ميلاد، أو كلمات تنتهي بلواحق معينة (مثل الكلمات التي تدل على الانفعالات).
- **أهم الرموز البسيطة في Regex:**
  - \d+: يبحث عن أي تسلسل من الأرقام (مثل درجات الطلاب).
  - \w+: يبحث عن أي كلمة.

- [a-z]+: يبحث عن كلمات باللغة الإنجليزية (مفيد في المصطلحات الأجنبية).

■ **مثال تطبيقي:** استخراج درجات الذكاء من نص تقرير:

Python

```
import re
report = "حصل الطالب على درجة 115 في اختبار الذكاء ودرجة 85 في التحصيل."
grades = re.findall(r'\d+', report)
print(grades)  # المخرجات: ['115', '85']
```

3. التعرف على الكيانات المسماة (Named Entity Recognition - NER)

هذه التقنية تسمح للحاسوب بـ تظليل الكلمات الهامة وتصنيفها آلياً. الحاسوب هنا يفهم أن أحمد هو **شخص (Person)**، وجامعة القاهرة هي **مؤسسة (Organization)**، و2024 هو **تاريخ (Date).**

في علم النفس التربوي: نستخدم NER لاستخراج أسماء الطلاب، المؤسسات التعليمية، والمواقع الجغرافية المذكورة في مئات التقارير لعمل خرائط إحصائية لمناطق انتشار صعوبات التعلم مثلاً.

المكتبات المستخدمة: مكتبة **spaCy** ومكتبة **NLTK** توفر نماذج جاهزة ومدربة للقيام بهذه المهمة.

4. تحليل الآراء والمشاعر (Sentiment Analysis) بشكل مبدئي

كأخصائيين، يهمنا معرفة النبرة في كتابات الطلاب أو المعلمين. تحليل المشاعر يساعدنا في تصنيف النصوص إلى (إيجابية، سلبية، محايدة) بناءً على الكلمات المفتاحية المكتشفة.

مثال: إذا كتب الطالب أشعر بإحباط شديد، سيقوم البرنامج برصد كلمة إحباط وتصنيف الحالة ضمن المشاعر السلبية.

5. أخلاقيات استخراج البيانات الشخصية

عندما نبدأ في استخراج الأسماء والتشخيصات آلياً، نصبح أمام مسؤولية أخلاقية كبرى:

إخفاء الهوية (Anonymization): قبل معالجة البيانات ونشرها، يجب استبدال الأسماء الحقيقية برموز (مثل الطالب أ، الطالب ب) لحماية الخصوصية.

الموافقة المستنيرة: يجب إبلاغ المشاركين (أو أولياء أمورهم) أن بياناتهم ستُعالج آلياً لأغراض بحثية.

تجنب التحيز: الخوارزميات قد تكون منحازة أحياناً؛ لذا يجب على الأخصائي مراجعة النتائج وعدم اعتمادها كحقيقة مطلقة دون فحص بشري.

الخلاصة والتطبيقات

لقد تعلمنا اليوم كيف نبدأ في الحفر داخل النصوص لاستخراج الكنوز المعلوماتية. استخدام Regex و NER يحولنا من قراء للنصوص إلى محللين للبيانات الضخمة.

■ **تمرين للتفكير:**

تخيل أن لديك ملفاً نصياً يحتوي على مئات الملاحظات السلوكية للمعلمين. كيف يمكن لتقنية استخراج الكيانات (NER) أن تساعدك في معرفة أي المعلمين أكثر رصداً للمشكلات السلوكية؟

في الدرس القادم (الجزء الثاني)، سنتعلم كيف نطور هذه المهارات لنبني قواعد بيانات حقيقية من هذه المعلومات المستخرجة.[3]