المحاضرة 12

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 11 كما ورد سابقاً...)

الدرس الثاني عشر: تطبيع النصوص في بايثون (Text Normalization)

يمثل **تطبيع النصوص ((Normalization** المرحلة الختامية في سلسلة معالجة البيانات النصية. بينما ركزنا في الدرس السابق على توحيد أشكال الحروف العربية (Unification)، فإن التطبيع يهدف إلى تحويل النص بالكامل إلى صيغة معيارية أو قياسية يسهل على الخوارزميات والبرامج الإحصائية التعامل معها دون لبس. بالنسبة لباحث علم النفس، التطبيع هو الضمانة الأكيدة لصحة القياس الرقمي للظواهر التربوية.

1. الفرق بين التوحيد والتطبيع

لتبسيط الأمر لطلبة علم النفس، يمكننا القول أن:

- **التوحيد:** هو معالجة أخطاء الكتابة (مثل: أ، إ ← ا).
- **التطبيع:** هو معالجة بنية النص وعناصره غير اللغوية، مثل الروابط، الإيموجي، تكرار الحروف غير المبرر، وتوحيد الأرقام.

2. عمليات التطبيع الأساسية في السياق التربوي

يتضمن تطبيع النصوص عدة خطوات تقنية حاسمة:

1. **تطبيع اليونيكود (:(Unicode Normalization** استخدام صيغ مثل NFKC لضمان أن الرموز المركبة تظهر بشكل موحد برمجياً، مما يمنع حدوث أخطاء خفية عند البحث في قاعدة بيانات

الطلاب.

2. **إزالة الضجيج الرقمي (Noise Removal):** حذف الروابط الإلكترونية (URLs)، عناوين البريد، والرموز الخاصة (#, @, *) التي قد توجد في استجابات الطلاب عبر المنصات التعليمية.
3. **معالجة الحروف المكررة:** الطلاب غالباً ما يعبرون عن مشاعرهم بتطويل الكلمات (مثل: حزيييييين أو راااائع). التطبيع يعيدها لأصلها (حزين، رائع) لضمان دقة تحليل المشاعر وعدم تشتت العد الإحصائي.
4. **توسيع الاختصارات:** في التقارير المكتوبة بالإنجليزية أو المترجمة، يتم تحويل الاختصارات إلى كلمات كاملة (مثل: u $\rightarrow$ you) لضمان فهم السياق.
5. **تطبيع الأرقام وعلامات الترقيم:** تحويل الأرقام الهندية (١، ٢) إلى أرقام عربية (1, 2) وتوحيد علامات الترقيم العربية (؟، ؛، ،) إلى نظائرها اللاتينية في حال كان البرنامج الإحصائي يتطلب ذلك.

3. التطبيق العملي: تنقية بيانات التنمر الإلكتروني

تخيل أنك تبحث في ظاهرة التنمر الإلكتروني في مدرسة ما، وجمعت تعليقات الطلاب من منصة مدرسية. النص الخام سيكون مليئاً بالإيموجي، الروابط، والكلمات الممدودة.

■ **مثال برمجي (التطبيع باستخدام Regex وبايثون):**

Python

```
import re

# نص خام من تعليق طالب
raw_comment = أنا أشعر بالاحباط الشديدددددد!!!! زوروا هذا الرابط http://help.me

# 1. إزالة الروابط
clean_text = re.sub(r'http\S+', '', raw_comment)

# 2. إزالة تكرار الحروف (أكثر من حرفين مكررين)
clean_text = re.sub(r'(.)\1+', r'\1', clean_text)

# 3. إزالة علامات الترقيم الزائدة
clean_text = re.sub(r'[!]+', '!', clean_text)

print(clean_text) # المخرجات: أنا أشعر بالاحباط الشديد! زوروا هذا الرابط
```

4. أهمية التطبيع في التحليل الاجتماعي الحسابي

من منظور علم النفس التربوي والاجتماعي، يساهم التطبيع في:

- **تحسين دقة التصنيف:** الخوارزميات التي تصنف الطلاب (مثلاً: معرض للفشل الدراسي) تعمل بدقة أعلى بنسبة تصل إلى 18% عند استخدام نصوص مطبعة.
- **تسهيل عنقودة النصوص (Clustering):** تجميع الحالات المتشابهة دلالياً تحت عنوان واحد (مثل: صعوبات تعلم) يصبح مستحيلاً بدون تطبيع الكلمات المفتاحية.
- **التمثيل الاتجاهي (Word Embeddings):** تحويل الكلمات إلى متجهات رياضية يتطلب نصاً نقياً جداً لفهم العلاقات الدلالية العميقة بين مفاهيم مثل الدافعية والتحصيل.

5. الاعتبارات الأخلاقية والمهنية الختامية

في نهاية هذا المقرر، يجب أن نتذكر دائماً أن معالجة النصوص ليست مجرد أكواد، بل هي أمانة مهنية:

1. **إخفاء الهوية (Anonymization):** كجزء من عملية التطبيع، يجب استبدال الأسماء الحقيقية برموز آمنة قبل المعالجة الضخمة لحماية خصوصية الطلاب.
2. **جودة البيانات وصحتها:** المؤسسة الجامعية تضمن جودة البيانات عبر الملاءمة والدقة والوضوح، وهي قيم يحققها التطبيع الرقمي.
3. **تجنب التحيز الخوارزمي:** يجب الحذر من أن تؤدي عملية التطبيع (مثل حذف الإيموجي) إلى فقدان نبرة الصوت أو المعنى النفسي الكامن وراء الانفعالات.

خاتمة المقرر:

يا أبنائي، لقد انتقلنا طوال هذه المحاضرات الاثنتي عشرة من فهم عتاد الحاسوب إلى بناء خوارزميات تفهم النفس البشرية عبر لغتها. إن الحاسوب في يد الأخصائي النفسي التربوي المعاصر ليس مجرد آلة كاتب، بل هو مكبر مجهري يرى ما لا تراه العين المجردة في آلاف الصفحات والتقارير.

لقد تعلمتم:

- كيف تنظمون بيئتكم الرقمية وتديرون ملفاتكم باحترافية.
- كيف تطوعون لغة بايثون لتحليل المشاعر واستخراج الكيانات.
- كيف تحولون النتائج إلى تقارير PDF وإكسال رسمية تخدم القرار التربوي.
- كيف تنقون وتوحدون وتطبعون النصوص العربية المعقدة لضمان صدق البحث العلمي.

التوصية الختامية:

التكنولوجيا تتطور، لكن المنهج العلمي يظل ثابتاً. استخدموا هذه الأدوات بذكاء ومسؤولية أخلاقية، واجعلوا من بايثون شريكاً لكم في فهم وتطوير المتعلم، فهدفنا الأسمى دائماً هو الإنسان.

تمنياتي لكم بمسيرة مهنية وبحثية مشرقة، عامرة بالإبداع والتميز الرقمي.