

دراسة في كلية تكنولوجيا المعلومات تبحث استخلاص الخصائص النصية وعنفقتها

بحثت دراسة في كلية تكنولوجيا المعلومات أعدتها الدكتورة غيداء عبد الحسين، والباحثة رشا نوري شعلان استخلاص الخصائص النصية وعنفقتها باعتماد التحليل الدلالي. أوضحت الدراسة أن التطور الحاصل مؤخراً في مجال التجارة الإلكترونية قد أدى إلى ظهور تحديات جديدة واجهت الباحثين في مجال معالجة آراء العملاء غالباً الأشخاص الذين يبحثون عن تجارة الأعمال يعتمدون على آراء زبائن آخرين لتكون مصدراً لهم في اتخاذ قرارات الشراء أو عقد الصفقات، ومع ذلك غرلة المواقع الكبيرة التي تحوي الكثير من الآراء عن خدمات معينة تعتبر مهمة شاقة وتستغرق وقتاً طويلاً، بالإضافة إلى الأبعاد الكثيرة في المساحة النصية لا تزال واحدة من الاهتمامات الرئيسية بسبب التعقيد الحسابي والدقة في تجميع النص. وتناولت الدراسة تحليل مجموعة من آراء العملاء الخاصة بالمطاعم تحليلاً ونحوياً ودلالياً لاستنتاج الخصائص النحوية والدلالية السياقية لاستثمار استخراج الخصائص النصية في تحديد التشابه، وتم اعتماد مجموعة من القواعد التي تم بناؤها لاستخراج الخصائص الدلالية لآراء العملاء التي تم اعتمادها في عملية تعدين (الأسماء المستخلصة من الآراء و توصيفاتها) وتقليل الأبعاد.

وتم استخلاص الخصائص الدلالية بثلاثة أساليب، وهي طريقة الظهور المشترك، وطرق التشابه المعتمدة على معجم البيانات، وطرق التشابه الموزعة لاستخراج خصائص دلالية للأسماء المستخلصة من الآراء وتوصيفاته، وتم استخدام خوارزمية تجميع لتجميع الأسماء المستخلصة التي لها قيم تشابه. وأظهرت النتائج كفاءة ومشجعة، بالإضافة إلى أن القواعد التي تم توليدها قادرة على استخلاص كل التوصيفات التابعة لأسم معين، وكذلك التجمعات الناتجة حصلت على قيم جيدة جداً لمقاييس الدقة والمراجعة، Yelp academic challenges، وتم تنفيذ النظام المقترح على بيانات مجهزة من الموقع، وتم تنفيذ النتائج باستخدام (Java platform on Net Beans IDE 8.0.2).

رافع عبد القادر