

Abstrakt

mgr Katarzyna Wójcik

tytuł rozprawy: Ocena miar podobieństwa dokumentów tekstowych na potrzeby automatycznej analizy opinii konsumenckich

Automatyczna analiza opinii konsumenckich to dziedzina badań, która może mieć znaczący wpływ na zarządzanie przedsiębiorstwem. Wielu konsumentów podejmuje decyzje po uprzednim przeszukaniu Internetu w poszukiwaniu opinii innych użytkowników. Z punktu widzenia zarządzania przedsiębiorstwem wiedza pochodząca z takich opinii jest istotna i może wpływać na podejmowane decyzje. Kadra zarządzająca dzięki automatycznej analizie opinii konsumenckich może w łatwy i szybki sposób uzyskać informacje o tym jak ona sama oraz jej produkty są postrzegane przez konsumentów. Do analizy opinii konsumentów, jako specyficznego rodzaju tekstów, wyrażających subiektywny stosunek autora wypowiedzi do jej przedmiotu, wykorzystywane są metody eksploracyjnej analizy tekstów, z uwzględnieniem tych bazujących na pomiarze podobieństwa lub odległości pomiędzy dokumentami tekstowymi.

Rozprawa składa się z czterech rozdziałów. Pierwszy z nich poświęcony jest zdefiniowaniu i scharakteryzowaniu opinii konsumentów w kontekście społeczeństwa informacyjnego. W rozdziale tym przybliżone są obszary nauki, dla których informacja jest przedmiotem badań oraz działalności ekonomicznej, w których informacje mają praktyczne zastosowanie. Następnie przytoczone są definicje danych, informacji i wiedzy jako kolejnych pojęć w modelu hierarchicznym informacji. Ostatnia część rozdziału poświęcona jest roli opinii konsumentów jako nośnika informacji w społeczeństwie informacyjnym, oraz jako przedmiotu badań z szeroko pojętego zarządzania.

Rozdział drugi skupiony jest wokół zagadnień dokumentów tekstowych oraz pomiaru podobieństwa pomiędzy nimi. Zdefiniowane jest tam pojęcie dokumentu tekstowego oraz pojęcia tekstu i pliku tekstowego. Następnie zdefiniowane i krótko scharakteryzowane są również obszary eksploracyjnej analizy tekstów oraz przetwarzania języka naturalnego. W dalszej części rozdziału zdefiniowane są pojęcia podobieństwa i odległości na gruncie analizy danych. Ostatnia część rozdziału poświęcona jest omówieniu różnych metod reprezentowania dokumentów tekstowych w systemach komputerowych oraz wyznaczania podobieństwa pomiędzy nimi w tym kontekście.

Trzeci rozdział rozprawy traktuje o automatycznej analizie opinii konsumentów. W tej części pracy omówiony jest proces automatycznej analizy opinii konsumentów wraz z różnymi podejściami i metodami badawcze stosowane przy jego realizacji. Wspomniane są tu sposoby uwzględniania wiedzy zewnętrznej w automatycznej analizie opinii konsumentów. Kolejna część rozdziału poświęcona jest przedstawieniu autorskiego modelu automatycznej analizy opinii konsumentów. Ostatnia część tego rozdziału dotyczy oceny przydatności miar podobieństwa dokumentów tekstowych w automatycznej analizie opinii konsumentów.

Czwarty, a zarazem ostatni rozdział rozprawy, to opis przeprowadzonych badań empirycznych. Znajdują się w nim zarówno opis procedury badawczej jak i omówienie uzyskanych wyników. Całość pracy kończą wnioski i weryfikacja realizacji celów rozprawy oraz postawionych w niej hipotez.