

ABSTRAK

Perkembangan teknologi informasi yang sangat pesat berdampak pada banyaknya ketersediaan informasi yang dapat dengan mudah diakses oleh siapapun dan di mana pun tempatnya. Namun, di sela-sela kemudahan tersebut, terkadang masyarakat juga mengalami kesulitan untuk memperoleh informasi yang sebenarnya mereka butuhkan. Oleh karena itu, dikembangkan sebuah sistem yang disebut *Question Answering System* (QAS), yaitu sebuah sistem yang dapat mengolah pertanyaan yang diberikan oleh pengguna dan kemudian dapat memberikan hasil atau jawaban berupa informasi yang diharapkan dapat sesuai dengan apa yang diinginkan oleh pengguna tersebut. Pengolahan tersebut tak terlepas dari *Natural Language Processing* (NLP) karena pertanyaan yang diberikan oleh pengguna merupakan bahasa alami yang harus diolah terlebih dahulu agar dapat diproses untuk melakukan Temu Balik Informasi (TBI) dan menghasilkan jawaban berupa informasi yang diinginkan pengguna. Untuk mengoptimasi proses QAS, dapat dilakukan klusterisasi dari pertanyaan yang diberikan pengguna. Hal tersebut dilakukan agar proses pencarian informasi dapat dilakukan hanya pada data-data yang berkaitan dengan pertanyaan tersebut. Proses klusterisasi ini kemudian dapat menentukan pertanyaan yang diberikan masuk ke dalam kluster mana, dan kemudian dapat dikelompokkan menurut aturan *5W2H*. *K-Means Clustering* merupakan salah satu algoritma yang paling efisien digunakan dalam proses klusterisasi karena dapat menghasilkan kluster-kluster yang terbagi dengan baik. Metode *Elbow* dapat mengoptimasi proses dari *K-Means Clustering* dengan menghitung terlebih dahulu nilai *Sum of Squared Errors* (SSE) dari tiap-tiap nilai kluster sehingga diperoleh jumlah kluster paling optimal. Hasil dari penelitian ini menunjukkan bahwa *K-Means Clustering* yang dioptimasi dengan metode *Elbow* dapat menghasilkan kluster-kluster yang sesuai dengan aturan *5W2H* dari data pertanyaan seputar otomotif, dibuktikan pula dengan hasil *Confusion Matrix* berupa nilai *precision* 90,85%, *recall* 88,67%, dan *f1-score* 88,75%. Diharapkan hasil dari penelitian ini dapat dikembangkan kembali untuk membangun sistem QAS dalam dunia otomotif.

Kata Kunci: *K-Means Clustering*, Metode *Elbow*, *Natural Language Processing*, Otomotif, *Question Answering System*