

Abstract

The number of verses in the Qur'an that is not small will be difficult and time consuming if done manually. Building a search system in the Qur'anic verse using the Indonesian Arabic-Latin equivalent will be very helpful for the Muslim community in Indonesia, especially for those who are not familiar with Arabic writing. In this study, a verse search system will be built on the Al-Qur'an based on phonetic similarity, more details about the handling of the cross verses in the Al-Qur'an. The system was built using the Jaro-Winkler algorithm to calculate the value of similarity and using the N-Grams algorithm for ranking documents. The same study has been done before with the name Lafzi +, with MAP 90% and 93% recall. In previous studies cases such as *nun wiqoyah* at the end of the verse could not be handled, so the system could not handle the search for the entire Qur'an. So to complete the previous research, in this study added rules other than pre-existing rules so that they can handle *nun wiqoyah* at the end of the verse. Rules are added to the phonetic coding process, for example in Surah An-Nisa verses 37 and 38, namely *فَخُورًا (۳۶) الَّذِينَ يَبْخُلُونَ* the verse contains *nun wiqoyah* at the end of the verse, in the Lafzi system + the verse produces a phonetic encoding "fahuranilazinayabhalun", by adding the rules of *nun wiqoyah* in this system result in the phonetic coding of the verse pieces becoming "pahuranxalajinayabhalun", to match the results of the phonetic coding stored in the database. By applying the Jaro-Winkler method to calculate the value of similarity and N-Grams for ranking documents and adding *nun wiqoyah* rules, this system generates 94% MAP and 92% recall. The results of this study indicate an increase in MAP, this shows that this system can improve the accuracy of systems that have been built before.

Keywords— al-quran, cross-verses, phonetic, jaro-winkler, n-gram.