

**INITIATION NATURAL LANGUAGE PROCESSING (NLP) AND
CLASSIFICATION OF QUESTIONS FOR MAKING
"STATBOT" (STATISTICS CHATTER BOT)**

**(Study Case: Information of Departement Statistics
Islamic University of Indonesia)**

Nita Tri Anggraini

*Department Statistics, Faculty of Mathematics and Natural Science
Islamic University of Indonesia*

ABSTRACT

Current technological developments. Where communication cannot be done officially, but it can be done using the internet. People today are wondering something that cannot be discussed on the internet. Interesting technology can make communication more interactive like interaction with humans, namely technology Chatter Bot or Chatbot. As in the Statistics Study Program, what often happens is key compilation, internship program, closing theory, examinations completed, and so on. Many students did not get this information, so that some students had to ask such things repeatedly to the department. Statistics UII, where chatbots will help or assist UII Statistics students specifically to get the information needed and help so that the UII Department of Statistics cannot be re-accessed information related to majors. Chatbot technology will be easier to use if you use a language that is used for everyday communication by humans, therefore Natural Language processing (NLP) is needed. Because this research data forms text, word weighting will be carried out using the Term Freq-Invers Document Frequency (TF-IDF) method and a Topic Modeling with Latent Dirichlet Allocation (LDA) is done to connect the words in the question into the question and see the interrelationship between the word. Then the question answers with Support Vector Machine (SVM). The results of the classification carried out obtained an evaluation of 0.888%. then the chatbot's ability to answer the questions raised is 70%.

Keywords : Chatbot, NLP, Topic Modelling, LDA, SVM