

Dyah Putri Tsaniyah, 2019. **Text Mining for Classification of Society's Comments about BPJS Kesehatan on *Twitter* with the Naive Bayes Classifier Method.** This final project is guided by Dr. Nur Chamidah, M.Si, and Drs. Suliyanto, M.Si, S-1 Study Program of Statistics, Department of Mathematics, Faculty of Science and Technology, Airlangga University, Surabaya

---

---

### ABSTRACT

Along with the current technological developments, the public can comment on the Indonesian government's program through Twitter's social network. Badan Penyelenggara Jaminan Sosial (BPJS) is one of the largest organizers of the health insurance program in Indonesia that often gets comments in the form of complaints or requests for information from the public. Based on data available until January 2019 the number of participants of the National Health Insurance program reached 215 million people, equivalent to 80% of the total population of Indonesia. This research was conducted with the aim to find out the steps in classifying the comments of Indonesian people by using algorithms and programs using the Naive Bayes Classifier method, and knowing the topics that are often discussed by Indonesian people related to BPJS Kesehatan. The data used are 1078 Indonesian-language comments taken from Twitter in February 2019 - March 2019 are divided into 80% training data and 20% testing data. Based on the data obtained results 37% of comments included in the category of membership with the word that appears most often is "cards", 45.6% of comments included in the service category with the words that appear most often are "hospital", and 17.4% comments included in the category of bills with the words that appear most often are "pay". The classification results in the training data get classification accuracy of 99%, and the testing data get results with classification accuracy of 98%.

**Keywords:** BPJS Kesehatan, *Classification, Naive Bayes Classifier*