

Deteksi Produk Obat Aborsi Secara Otomatis Studi Kasus PT XYZ = Abortion Drug Product Detection Automatically Study Case PT XYZ

Muhammad Faisal M, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20528320&lokasi=lokal>

Abstrak

Terdapat beberapa media online yang ditutup oleh Kementerian Komunikasi dan Informatika (Kemkominfo) dikarenakan menjual obat aborsi. Hal tersebut karena aborsi merupakan tindakan yang dilarang yang tertulis pada Kitab Undang-Undang Hukum Pidana (KUHP) pada pasal 346. Oleh karena itu, agar situs dari PT XYZ tidak ditutup oleh Kemkominfo, PT XYZ melakukan penanganan terkait peredaran obat aborsi ini yaitu dengan pending system. Namun, pending system hanya mendeteksi judul dari produk dengan menggunakan kata kunci spesifik yang berhubungan dengan obat aborsi yang diinput oleh tim sehingga masih terdapat produk obat aborsi yang lolos beredar karena terdapat produk yang menggunakan kata kunci yang umum dan gaming keyword. Oleh karena itu, penelitian ini membahas terkait penerapan text mining untuk membangun sebuah classification model yang berasal dari korpus obat aborsi yang ada di PT XYZ yang akan digunakan untuk pendeteksian obat aborsi kedepannya yang ada di PT XYZ.

Penelitian ini menggunakan model CRISP-DM untuk siklus hidup data mining. Selain itu, untuk membangun suatu classification model, Penelitian ini melakukan percobaan terhadap dua algoritme diantaranya adalah Naive Bayes dan Support Vector Machine dengan metode k-fold cross validation. Selain itu, penelitian ini menggunakan data harga sebagai fitur tambahan dari model yang dibangun. Untuk penentuan classification yang terbaik dilakukan evaluasi performa dari setiap classification model dengan menggunakan confusion matrix dengan parameter accuracy, recall, precision, f1-measure, dan AUC. Penelitian ini menggunakan beberapa kriteria dalam penghapusan duplikasi data untuk menghindari data bias. Model terbaik yang didapatkan yaitu model SVM dengan fitur harga yang memiliki nilai accuracy 99.82%, f1-score 99.79%, dan AUC 99.98%. Hasil dari model yang telah dianalisis pada penelitian ini dapat digunakan oleh PT XYZ untuk mendeteksi produk obat aborsi agar mengurangi kesempatan penjual menjual produk obat aborsi yang di PT XYZ. Selain itu, penelitian ini dapat memberikan gambaran untuk penelitian akademis berikutnya terkait keseluruhan proses dari text mining.

.....There are several online media that were closed by the Ministry of Communication and Information (Kemkominfo) due to selling abortion drugs. This is because abortion is a prohibited act which is written in the Criminal Code (KUHP) in article 346. Therefore, in order PT XYZ is not closed by the Ministry of Communication and Information, PT XYZ create system that try to handle the circulation of abortion drugs, namely pending system. However, the pending system only detects the title of the product by using specific keywords related to abortion drugs that are inputted by the team so that there are still abortion drug products that pass through the system because there are products that use general keywords and gaming keywords. Therefore, this study discusses the application of text mining to build a classification model derived from the abortion drug corpus at PT XYZ which will be used for the detection of abortion drugs in the future at PT XYZ.

This study uses the CRISP-DM model for the data mining life cycle. In addition, to build a classification model, this study conducted experiments on two algorithms including Naive Bayes and Support Vector

Machine with the k-fold cross validation method. In addition, this study uses price data as an additional feature of the built model. To determine the best classification, the performance evaluation of each classification model is carried out using a confusion matrix with parameters accuracy, recall, precision, f1-measure, and AUC. This study uses several criteria in eliminating duplication of data to avoid data bias. The best model obtained is the SVM model with a price feature that has an accuracy value of 99.82%, f1-score 99.79%, and AUC 99.98%. The results of the model that had been analyzed in this study can be used by PT XYZ to detect abortion drug products in order to reduce the chance for sellers to sell abortion drug products at PT XYZ. In addition, this research can provide an overview for the next academic research related to text mining process.