

---

# Penerapan *Question Entailment* pada Sistem Tanya Jawab Topik Kesehatan Anak

Arya Prima Al Aftar<sup>1</sup>, Ade Romadhony<sup>2</sup>, Hasmawati<sup>3</sup>

<sup>1,2,3</sup>Fakultas Informatika, Universitas Telkom, Bandung

<sup>1</sup>aryaprima@students.telkomuniversity.ac.id, <sup>2</sup>aderomadhony@telkomuniversity.ac.id,

<sup>3</sup>hasmawati@telkomuniversity.ac.id

---

## Abstrak

Kesehatan merupakan salah satu hal yang paling penting dalam kehidupan manusia khususnya pada anak-anak, yang memungkinkan anak-anak dapat tumbuh dan berkembang secara baik. Untuk menjaga kesehatan maka orang tua juga harus memiliki informasi yang tepat mengenai cara menjaga kesehatan, pola makan, pola hidup yang baik. Dalam mendapatkan informasi itu dibutuhkan sebuah platform untuk memudahkan pencarian informasi mengenai kesehatan terutama kesehatan anak. Pada penelitian ini diterapkan sebuah sistem tanya jawab atau question answering dengan menerapkan Question Entailment untuk memudahkan mencari pertanyaan mengenai kesehatan anak dan mendapatkan jawaban yang tepat. Dataset yang digunakan dibuat berdasarkan hasil pengumpulan daftar pertanyaan dan jawaban dari buku tanya jawab kesehatan anak dan FAQ di internet. Model Question Entailment dibangun berdasarkan training korpus yang dibuat dan diuji menggunakan algoritma Support Vector Machine(SVM), Logistic Regression, Naïve Bayes, dan J48 Decision Tree. Hasil pengujian menunjukkan bahwa algoritma SVM memberikan performa terbaik dalam mengidentifikasi pertanyaan serupa, dengan metrik precision, recall, dan f1(f-measure).

**Kata kunci:** Kesehatan anak, Question Answering, Question Entailment, Training korpus, Support Vector Machine (SVM)

---

## Abstract

Health is one of the essential things in human life, especially in children, which allows children to grow and develop properly. Parents must also have the correct information about maintaining health, diet, and an excellent lifestyle. In getting that information, a platform is needed to facilitate the search for health information, especially children's health. In this research, a Question Answering System was implemented by applying Question Entailment to make it easier to find questions about children's health and get the correct answers. The dataset is created based on collecting a list of questions and answers from a child health question and answer book and Internet FAQs. The Question Entailment model is built based on a training corpus that has been created and tested using Support Vector Machine (SVM), Logistic Regression, Naïve Bayes, and J48 Decision Tree algorithm. The experimental results show that the SVM method gives the best performance on identifying similar question, with precision, recall, and f1(f-measure) metrics.

**Keywords:** Children's health, Question Answering, Question Entailment, Training corpus, Support Vector Machine (SVM)

---

## 1. Pendahuluan

### Latar Belakang

Kesehatan merupakan salah satu hal yang penting bagi manusia, yang memungkinkan untuk setiap orang dapat hidup produktif secara sosial dan ekonomis, namun sejak *Coronavirus* datang ke Indonesia pada Maret 2020 [1], angka kesehatan warga Indonesia menjadi menurun dan rentan terkena *Coronavirus* ataupun penyakit lainnya, sehingga masyarakat Indonesia harus lebih menjaga kesehatan serta kebersihan tubuhnya. Selama pandemi COVID-19 ini masyarakat Indonesia dianjurkan untuk melakukan aktifitas hanya dirumah saja dan tidak dianjurkan untuk keluar rumah jika tidak memiliki kegiatan yang penting terutama pada anak-anak tidak dianjurkan untuk sering ke fasilitas kesehatan terutama anak dengan usia yang masih belia karena belum memiliki imun yang kuat dan juga program dari Posyandu mungkin terhambat pelaksanaannya sehingga komunikasi orang tua bila memiliki keluhan atau pertanyaan terkait anak-anak nya sebaiknya dapat