



DIGITAL  
TALENT  
SCHOLARSHIP

TA  
Thematic  
Academy

```
PORTA
; Configure Port B<7-0> as output
tristate_out 0

idle = 0;
if (data_read == 0) {
    data &= 0xf;
    if (data == 12) {
        if (digits == 4 && /* check code */
            digit1 == CODE1 &&
            digit2 == CODE2 &&
            digit3 == CODE3 &&
            digit4 == CODE4) {
            /* set B<4> if correct code */
            delay_sec(2);
            /* reset */
            digits = 0;
            data_read = 0;
        } else {
            digits++;
            data_read = 1;
        }
    }
}
```

# Data Science Business Understanding Analisis Topik Vaksinasi Covid-19 di Indonesia

Thematic Academy  
Digital Talent Scholarship  
Tahun 2021

## Daftar Isi Proposal Bisnis Data Science

<b>1</b>	<b>3</b>
<b>1.1</b>	<b>3</b>
<b>1.2</b>	<b>3</b>
<b>1.3</b>	<b>3</b>
<b>2</b>	<b>4</b>
<b>2.1</b>	<b>4</b>
<b>2.2</b>	<b>4</b>
<b>3</b>	<b>4</b>
<b>3.1</b>	<b>4</b>
<b>3.2</b>	<b>5</b>
<b>3.3</b>	<b>6</b>
<b>4</b>	<b>6</b>

### **1.**

## 2. Kelompok 1

- Nisa Hanum Harani
- Muhammad Rajab Fachrizal

### 1 Definisi Permasalahan

#### 1.1 Latar Belakang

Program vaksinasi pemerintah merupakan hal krusial untuk mengendalikan penyebaran virus corona di Indonesia. Respon masyarakat terhadap program vaksinasi ini, khususnya pada *platform* twitter, pun beragam. Ada yang mendukung, namun tidak sedikit yang menolak. Dinamika ini bisa dipengaruhi banyak hal, salah satunya persepsi masyarakat / diskursus sosial terkait vaksinasi. Hal ini tentunya mempengaruhi efektifitas pemerataan program vaksinasi di Indonesia. Untuk itu, diskursus masyarakat akan coba dipetakan sehingga terlihat jelas kecenderungan arah persepsi masyarakat terhadap program vaksinasi Covid-19. Besaranya jumlah data dan data yang bervariasi akan menyulitkan untuk mengekstraksi setiap topik tersebut jika dilakukan dengan proses manual. Oleh karena itu, proses mengekstraksi dan mempelajari pola yang tersembunyi dari setiap topik yang dibahas diperlukan suatu metode dan algoritma *machine learning* yang dapat memberikan hasil yang cepat dan tingkat akurasi yang terukur. Setelah mengekstrak dan membersihkan informasi, maka dilakukan label (positif, negative dan netral) . Data yang sudah dilakukan labelisasi kemudian akan di training dan dilakukan sentiment analisis mengani Vaksin yang beredar di Indonesia.

#### 1.2 Identifikasi Permasalahan Bisnis

Berdasarkan latar belakang yang telah dipaparkan sebelumnya, terdapat beberapa permasalahan, yaitu :

1. Bagaimana memetakan arah persepsi masyarakat terhadap program vaksin Covid-19 berdasarkan data pada *platform* twitter
2. Bagaimana kecenderungan arah persepsi masyarakat (positif, negative, netral) terhadap program vaksin Covid-19

#### 1.3 Asumsi, Kebutuhan, dan Limitasi

Berdasarkan identifikasi masalah yang telah dijabarkan maka kebutuhan dalam penelitian ini yaitu:

1. Akuisisi data twitter dengan kunci setiap vaksin yang dipakai di indonesia.
2. Hanya melakukan analisis twitter berbahasa indonesia.

## 2 Tujuan Proyek

### 2.1 Tujuan Bisnis

Analisis topik terkait vaksinasi Covid-19 di Indonesia ini bertujuan untuk mengetahui gambaran arah persepsi masyarakat (positif, negative, netral) terhadap program vaksinasi ini, khususnya pada *platform* twitter.

### 2.2 Measureable Outcomes / Business Criteria

Dapat memperoleh deskripsi topik pada data twitter terkait perkembangan isu terkini seputaran vaksin-vaksin yang telah digunakan oleh masyarakat di Indonesia. Hasil deskripsi topik juga dapat dimanfaatkan oleh peneliti untuk melanjutkan ke dalam proses opini mining.

## 3 Solusi dari Perspektif Bisnis

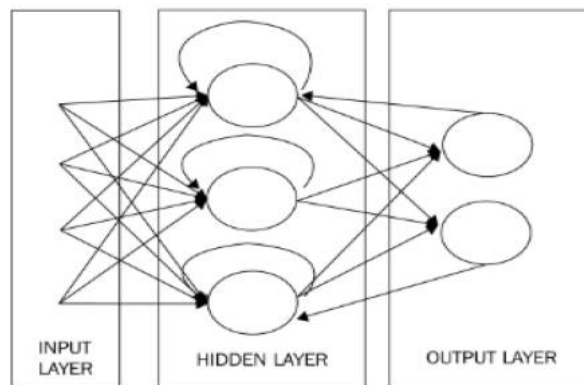
### 3.1 Solutions: Regression, Classification, Clustering, Recommendation, etc

Dalam studi kasus ini, untuk dapat menganalisis arah persepsi berdasarkan opini atau lebih dikenal sebagai sentiment analysis. Sentimen dapat diartikan sebagai pendapat ataupun pandangan yang didasarkan pada perasaan berlebihan terhadap sesuatu. Sentimen biasanya terdapat dalam pernyataan serta kalimat yang memiliki pendapat. Sentimen juga berguna untuk mengetahui perasaan yang diberikan oleh seseorang terhadap topik atau objek tertentu.

Analisis sentimen atau bisa juga disebut dengan *opinion mining* adalah studi komputasional dari opini orang lain, appraisal, serta emosi yang terdapat dalam entitas, *event*, dan atribut yang dimiliki. Tugas yang dilakukan oleh analisis sentimen yaitu mengelompokkan polaritas yang terdapat pada suatu teks, baik yang terdapat dalam dokumen, kalimat, maupun fitur tingkat aspek apakah pendapat yang dikemukakan bersifat positif atau negatif. Terdapat beberapa proses yang dapat dilakukan dalam *opinion mining* dan dibagi ke dalam beberapa tingkatan. Suatu sentimen atau pendapat dapat ditemukan pada *document level*, *sentence level*, maupun *entity level*.

*Recurrent Neural Network* (RNN) merupakan salah satu metode *deep learning* yang memiliki tugas untuk melakukan ekstraksi fitur dari suatu data secara lebih rinci. Terdapat beberapa

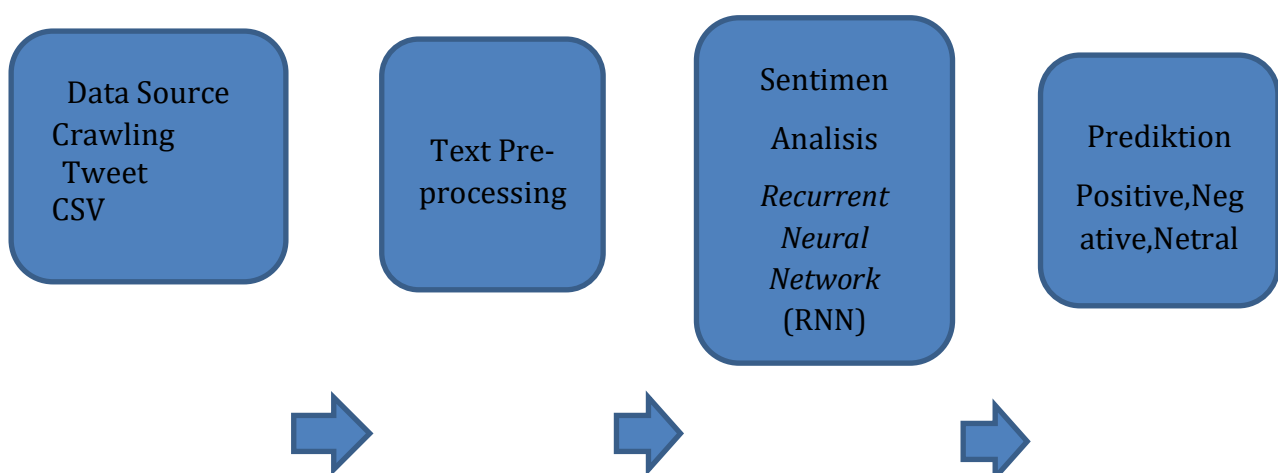
varian dari *Artificial Neural Network* (ANN), beberapa varian tersebut diantaranya dikelompokkan berdasarkan jumlah lapisan tersembunyi dan aliran data salahsatunya adalah RNN, dimana koneksi antar neuron dapat membentuk sebuah siklus. RNN adalah varian ANN yang memiliki fitur koneksi antara lapisan tersembunyi yang disebarkan melalui waktu untuk mempelajari urutan dan untuk melakukan pemrosesan data sekuensial. Selain itu RNN juga memiliki kemampuan yang sama dengan *convolutional network* yaitu dapat dengan mudah melakukan *scaling* terhadap gambar dengan memiliki dimensi yang besar, namun proses *scaling* tersebut membutuhkan waktu yang lebih lama dari biasanya untuk sebuah jaringan yang bukan spesialisasi pemroses data sekuensial.



Gambar 1 RNN

### 3.2 Analytical Approach: Data Cleansing, Data Preparation

Data yang digunakan bersifat tidak terstruktur dalam bentuk teks yang diperoleh dengan cara *crawling* pada *platform* twitter dengan kata kunci yang berkaitan dengan vaksinasi Covid-19 di Indonesia. Agar data dapat digunakan pada algoritma *machine learning*, data dalam bentuk teks tersebut kemudian dilakukan Text Pre-Processing untuk menghilangkan bagian atau teks yang tidak diperlukan sehingga mendapatkan data yang berkualitas.



### *3.3 Analytical Approach: Model Selection and Metric Analysis*

## **4 Instrumen Pengukuran Keberhasilan**

Hasil pemodelan topik modeling pada rentang waktu tertentu dapat membentuk topik dengan informasi atau deskripsi topik untuk setiap vaksin yang dipakai. Deskripsi opini adalah kecenderungan informasi yang muncul jika menggunakan vaksin tertentu, hasil ini dapat merangkum tanggapan masyarakat mengenai penggunaan vaksin di indonesia.



KOMINFO

Badan Penelitian dan Pengembangan SDM  
Kementerian Komunikasi dan Informatika