

Laporan Tugas 2

Kecerdasan Artificial - INF311

Diky Wahyudi - 2108107010031

Jenis Kasus

Pada tugas 2 Kecerdasan Artificial ini diangkat sebuah kasus untuk melakukan klasifikasi terhadap sebuah jenis berita. Dengan mengembangkan sebuah model untuk melakukan klasifikasi jenis berita dapat membantu dalam mengetahui jenis dari sebuah berita

Data yang akan digunakan adalah data dari sebuah portal berita yang telah diberikan label.

Dataset yang digunakan

Dataset yang digunakan adalah sebuah dataset yang berisikan berita dari sebuah portal berita di Indonesia dan pada dataset telah memiliki label untuk setiap berita.

Dataset dapat diakses melalui website Huggingface dengan link berikut:

<https://huggingface.co/datasets/jakartaresearch/indonews>

Jumlah Fitur

Pada dataset hanya terdapat 1 fitur yaitu isi dari berita

text
string · lengths



Liputan6.com, London - Lee Dixon khawatir Arsenal tak bisa merekrut Denis Suarez secara permanen musim panas nanti. Pasalnya sejauh ini, Suarez hanya menjadi cadangan bagi The...

Liputan6.com, Jakarta - Kasus dugaan penganiayaan terhadap dua remaja yang melibatkan Bahar bin Smith telah rampung dan dinyatakan lengkap atau P21 oleh Kejaksaan Negeri Cibinong, Senin, 4...

Liputan6.com, Jakarta Menanggapi aksi eks pekerja PT Freeport Indonesia(PTFI) yang menggelar tenda dan tidur di depan Istana Merdeka, Jakarta Pusat dan beberapa aksi sebelumnya, juru...

Liputan6.com, Medan - Sebanyak 81 kendaraan 4x4 dan 74 motor telah tiba seluruhnya di Medan untuk bersiap start Indonesia Off-road eXpedition 2019 hari Sabtu ini (9/2), dari Medan menuju...

Liputan6.com, Jakarta Indonesia akan melawan Thailand di Final Piala AFF U-22 pada Selasa 26 Februari 2019 pukul 18.30 WIB.

Liputan6.com, Jakarta - Komisi Pemberantasan Korupsi (KPK) memperpanjang penahanan Wakil Ketua reaktif DPR Taufik Kurniawan (TK) dalam kasus dugaan suap dan seleksi korupsi (DAK) Kemen...

Jumlah Label

Jumlah Label pada dataset yang digunakan adalah 5 kategori yaitu, bola, news, bisnis, tekno dan otomotif

```
[57] df["label"].unique()  
  
array(['bola', 'news', 'bisnis', 'tekno', 'otomotif'], dtype=object)
```

label
string classes

5 values
bola
news
bisnis
bola
bola
news

Jenis Neural Network yang Digunakan

Jenis neural network yang digunakan adalah ANN (Artificial Neural Network). Pada model Neural Network yang digunakan , terdapat 3 layer.

- Embedding Layer
- Flatten Layer
- Dense Layer

```
model = models.Sequential([  
    layers.Embedding(MAX_WORDS, 8, input_length=MAXLEN), # Output layer memiliki 8 dimensi  
    layers.Flatten(),  
    layers.Dense(5, activation='softmax') # Memiliki 5 output sesuai dengan jumlah jenis label  
)
```

Jenis Optimasi

Jenis optimisasi yang digunakan adalah RMSprop. RMSprop adalah singkatan dari "Root Mean Squared Propagation". RMSprop adalah algoritma optimasi yang digunakan untuk melatih neural network dengan berbasis gradient descent.

$$MeanSquare(w, t) = 0.9 MeanSquare(w, t-1) + 0.1 \left(\frac{\partial E}{\partial w}(t) \right)^2$$

Jenis Fungsi Aktivasi yang Digunakan

Fungsi aktivasi yang digunakan pada layer ke 3 adalah Fungsi Softmax.

$$\sigma(\vec{z})_i = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}}$$

```
model = models.Sequential([  
    layers.Embedding(MAX_WORDS, 8, input_length=MAXLEN), # Output layer memiliki 8 dimensi  
    layers.Flatten(),  
    layers.Dense(5, activation='softmax') # Memiliki 5 output sesuai dengan jumlah jenis label  
])
```


Jumlah Hidden Layer

Pada model yang dikembangkan hanya terdapat 1 hidden layer yaitu adalah Layer Flatten. Layer ini berfungsi untuk merubah data multi-dimensi menjadi 1 dimensi.

Dapat dilihat pada gambar disamping, hasil output pada layer pertama memiliki 2 dimensi kemudian pada layer Flatten akan dijadikan 1 dimensi.

$$750 \times 8 = 6000$$

```
[54] # Summary dari setiap layer  
model.summary()
```

Model: "sequential"

Layer (type)	Output Shape	Param #
embedding (Embedding)	(None, 750, 8)	320000
flatten (Flatten)	(None, 6000)	0
dense (Dense)	(None, 5)	30005

```
=====  
Total params: 350005 (1.34 MB)  
Trainable params: 350005 (1.34 MB)  
Non-trainable params: 0 (0.00 Byte)
```

Jumlah Total Hidden Node per Layer

Berdasarkan hasil disamping dapat dilihat jumlah nodes pada setiap layer

- Embedding Layer (750 Nodes)
- Flatten Layer (1 Nodes)
- Dense (5 Node)

Pada Flatten Layer hanya memiliki 1 node karena, pada layer ini hanya merubah data yang multi-dimensi menjadi data 1 dimensi.

Model: "sequential"

Layer (type)	Output Shape	Param #
embedding (Embedding)	(None, 750, 8)	320000
flatten (Flatten)	(None, 6000)	0
dense (Dense)	(None, 5)	30005

=====
Total params: 350005 (1.34 MB)
Trainable params: 350005 (1.34 MB)
Non-trainable params: 0 (0.00 Byte)

Jumlah Total Bobot

Berdasarkan hasil disamping dapat dilihat jumlah bobot yang ada pada setiap layer

- Embedding Layer (320.000 Bobot)
- Flatten Layer (0 Bobot)
- Dense Layer (30000 Bobot + 5 Bias)

Model: "sequential"

Layer (type)	Output Shape	Param #
embedding (Embedding)	(None, 750, 8)	320000
flatten (Flatten)	(None, 6000)	0
dense (Dense)	(None, 5)	30005

=====
Total params: 350005 (1.34 MB)
Trainable params: 350005 (1.34 MB)
Non-trainable params: 0 (0.00 Byte)

Thank You - Diky Wahyudi