

**Laporan Proyek Kecerdasan Buatan**  
**Zoo Animal Classification**



**ANGGOTA KELOMPOK :**

<b>Francisco Allenxeon</b>	- C14190118
<b>Christian Willson</b>	- C14190178
<b>Richardo Jason</b>	- C14190181
<b>Richson Sedjie</b>	- C14190185

## Pendahuluan

Proyek kami membuat Zoo Animal Classification dengan menggunakan Neural Network. Dimana dengan dataset yang kami miliki akan digunakan untuk melatih program yang ada untuk melakukan klasifikasi binatang sesuai dengan ciri-ciri yang dimiliki. Ciri-ciri yang dimiliki pun beragam seperti jumlah kaki yang dimiliki, memiliki tulang belakang atau tidak, dan ciri lainnya.

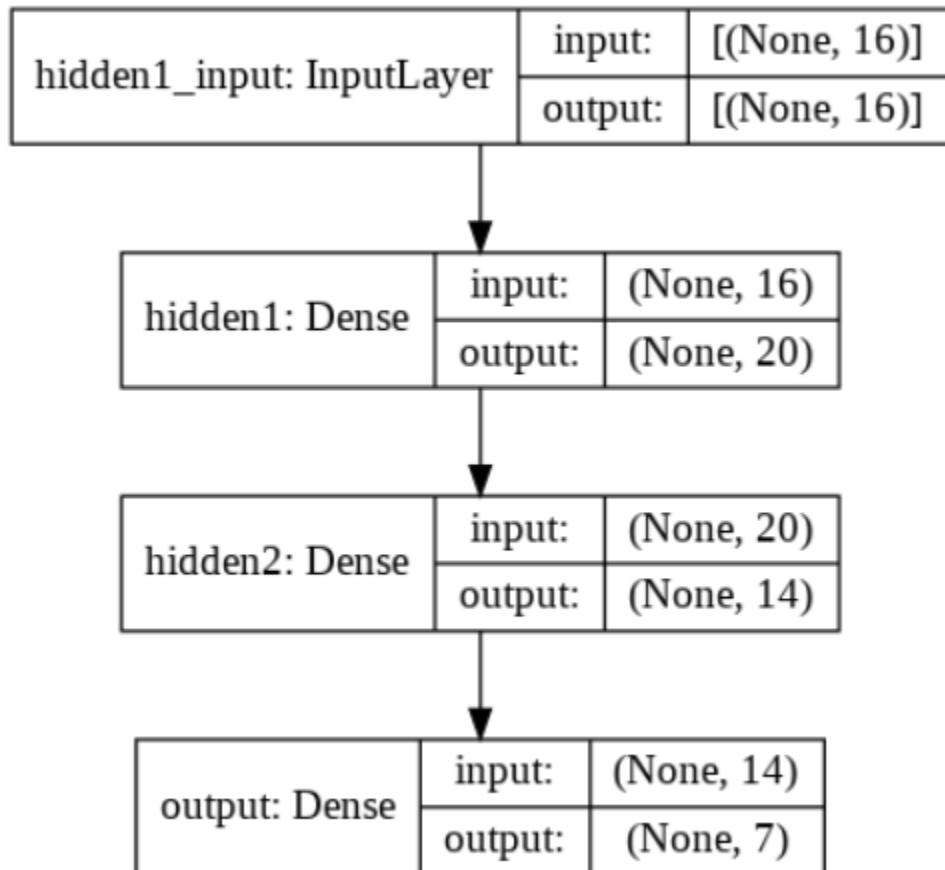
## Dataset

Dataset : <https://www.kaggle.com/uciml/zoo-animal-classification>

Alasan menggunakan Dataset :

Karena dataset yang kami pilih memiliki data yang cukup mudah diproses dan juga memiliki attribute yang cukup kompleks. Dimana dengan data tersebut, dapat digunakan untuk membuat suatu model yang bisa di train.

## Model



Perbandingan penggunaan data yang telah dinormalisasi dengan tidak

No	Accuracy	
	Without Normalization	MinMax Scaler
1.	0.9767	0.9302
2.	0.9535	0.9767
3.	0.9302	0.9534
4.	0.9535	0.9767
5.	0.9535	0.9534
6.	0.9535	0.9534
7.	0.9767	0.9767
8.	0.9535	0.9767
9.	0.9767	0.9767
10.	0.9767	0.9534
Avg	0.96045	0.96273

Bisa disimpulkan hasil dari 10 kali percobaan, kelompok kami menyimpulkan tidak terjadi perbedaan yang terlalu signifikan dikarenakan atribut yang dinormalisasikan hanya 1 sehingga tidak terlalu memberikan efek kepada penghitungan.

### Normalize Data

Untuk menormalisasi data, kami menggunakan MinMaxScaler(). Kelompok kami karena data kami memiliki batasan yang jelas dan juga tidak menggunakan distribusi normal sehingga kami menggunakan MinMaxScaler serta menggunakan MinMaxScaler karena outputnya memiliki batasan 0 - 1.

Untuk cara kerja dari MinMaxScaler() adalah

$$x_{scaled} = \frac{x - x_{min}}{x_{max} - x_{min}}$$

Semisal data suatu hewan memiliki kaki 4.

Jika di normalize maka akan  $X_{scaled} = 4-0/8-0 = \frac{1}{2} = 0.5$

Min = 0 karena ada hewan yang tidak memiliki kaki.

Max = 8 karena data kaki dengan value terbesar adalah 8 sesuai dengan dataset yang ada.

## Cara Kerja One-Hot Encoding

One-Hot Encoding bekerja dengan cara mengubah data menjadi angka biner yaitu 0 dan 1 . One-Hot Encoding tidak menormalize data yang digunakan tetapi hanya lebih ke arah mempersiapkan data yang akan digunakan tersebut. Cara kerjanya adalah dengan

Sebelum di One-Hot Encoding :

```
|array([1, 1, 4, 1, 1, 1, 4, 4, 1, 1, 2, 4, 7, 7, 7, 2, 1, 4, 1, 2, 2,
       1, 2, 6, 5, 5, 1, 1, 1, 6, 1, 1, 2, 4, 1, 1, 2, 4, 6, 6, 2, 6, 2,
       1, 1, 7, 1, 1, 1, 1, 6, 5, 7, 1, 1, 2, 2, 2, 2, 4, 4, 3, 1, 1, 1,
       1, 1, 1, 1, 1, 2, 7, 4, 1, 1, 3, 7, 2, 2, 3, 7, 4, 2, 1, 7, 4, 2,
       6, 5, 3, 3, 4, 1, 1, 2, 1, 6, 1, 7, 2])
```

\*Data Y merupakan data Class Type dari dataset zoo.csv

Sesudah dilakukan One-Hot Encoding :

```
C> array([[1., 0., 0., 0., 0., 0., 0.],  
[1., 0., 0., 0., 0., 0., 0.],  
[0., 0., 0., 1., 0., 0., 0.],  
[1., 0., 0., 0., 0., 0., 0.],  
[1., 0., 0., 0., 0., 0., 0.],  
[1., 0., 0., 0., 0., 0., 0.],  
[1., 0., 0., 0., 0., 0., 0.],  
[1., 0., 0., 0., 0., 0., 0.],  
[0., 0., 0., 1., 0., 0., 0.],  
[0., 0., 0., 1., 0., 0., 0.],  
[1., 0., 0., 0., 0., 0., 0.],  
[0., 1., 0., 0., 0., 0., 0.],  
[0., 0., 0., 1., 0., 0., 0.],  
[0., 0., 0., 0., 0., 0., 1.],  
[0., 0., 0., 0., 0., 0., 1.],  
[0., 0., 0., 0., 0., 0., 1.],  
[0., 1., 0., 0., 0., 0., 0.],  
[1., 0., 0., 0., 0., 0., 0.],  
[0., 0., 0., 1., 0., 0., 0.],  
[1., 0., 0., 0., 0., 0., 0.],  
[0., 1., 0., 0., 0., 0., 0.],  
[0., 0., 0., 0., 1., 0., 0.],  
[0., 0., 0., 0., 0., 1., 0.],  
[0., 0., 0., 0., 0., 0., 1.],  
[1., 0., 0., 0., 0., 0., 0.],  
[1., 0., 0., 0., 0., 0., 0.],  
[0., 1., 0., 0., 0., 0., 0.],  
[1., 0., 0., 0., 0., 0., 0.],  
[0., 1., 0., 0., 0., 0., 0.],  
[0., 0., 0., 0., 0., 1., 0.],  
[0., 0., 0., 0., 1., 0., 0.],  
[0., 0., 0., 0., 0., 1., 0.],  
[1., 0., 0., 0., 0., 0., 0.],  
[1., 0., 0., 0., 0., 0., 0.],  
[0., 0., 0., 0., 0., 0., 1.],  
[0., 0., 0., 0., 0., 1., 0.],  
[1., 0., 0., 0., 0., 0., 0.],  
[1., 0., 0., 0., 0., 0., 0.],  
[0., 0., 0., 0., 0., 0., 1.],  
[0., 0., 0., 0., 0., 1., 0.],  
[1., 0., 0., 0., 0., 0., 0.],  
[1., 0., 0., 0., 0., 0., 0.],  
[0., 0., 0., 0., 0., 0., 1.]])
```

One-Hot Encoding mengubah data class\_type yang awalnya berupa angka dari range 1-7 menjadi representasi biner dimana setiap posisi dalam representasi biner tersebut merepresentasikan class\_type dari binatang yang berkorespondensi dengan baris biner tersebut.

Seperti bisa dilihat dari Gambar pertama, bisa kita lihat class typenya adalah 1 dan pada gambar kedua kolom pertama sesuai dengan gambar pertama. Sehingga One Hot Encoding merubah datanya menjadi bisa dibilang seperti gambar dibawah.

	1(Mammal)	2(Bird)	3(Reptile)	4(Fish)	5(Amphibian)	6(Bug)	7(Vertebrate)
data1	1	0	0	0	0	0	0
data2	1	0	0	0	0	0	0
daata3	0	0	0	1	0	0	0

### Pemilihan Units

No	<b>12,10,7</b>	<b>20,14,7</b>	<b>64,32,7</b>
1.	0.9302	0.9302	0.9767
2.	0.8837	0.9767	0.9767
3.	0.9534	0.9534	0.9767
4.	0.9069	0.9767	0.9767
5.	0.9302	0.9534	0.9767
6.	0.9302	0.9534	0.9767
7.	0.9069	0.9767	0.9767
8.	0.8837	0.9767	0.9767
9.	0.9302	0.9767	0.9767
10.	0.9069	0.9534	0.9767
Avg	0.91623	0.96273	0.9767

Kesimpulan yang kelompok kami ambil adalah sama dengan apa yang ibu jelaskan dikelas. Pemilihan units yang semakin mendekati maka akan semakin bagus hasil yang diberikan. Dimana pengertian bagus tidak selamanya ditunjukkan dengan angka yang tinggi melainkan bisa dilihat juga dari variasi nya.

### Pemilihan Epoch & Batch Size

No	Epoch : 2 batch_size : 86	Epoch : 40 batch_size : 10	Epoch : 86 batch_size : 2
1	0.1245	0.9534	0.9809
2	0.0700	0.9773	0.9753
3	0.1128	0.9741	0.9882
4	0.1869	0.9330	0.9907
5	0.1050	0.9686	0.9860

<b>Avg</b>	0.11984	0.96128	0.98422
------------	---------	---------	---------

Kesimpulan yang kelompok kami ambil adalah semakin banyak epoch maka accuracy yang diberikan bisa semakin tinggi tetapi apabila epoch nya juga terlalu banyak maka akan menyebabkan *overfitting*. Batch size akan mempengaruhi jumlah data yang akan diiterasikan setiap epoch.

## Testing Inputan

Langkah-langkah melakukan testing data inputan :

1. Load data / file yang ingin di testing.
  2. Mengambil attribute yang diperlukan seperti kolom 1 hingga 16
  3. Memasukkan class type ke suatu variabel.
  4. Melakukan prediction melalui model yang telah di train.
  5. Kemudian membandingkan prediction dengan variabel yang berisi class type yang benar.

Prediction	Correct
0	0
1	5
2	2
3	0
4	0

Mammalia

Bug

Reptile

Mammalia

Mammalia

done

## **Kendala**

1. Pembuatan model yang menerima inputan user atau membaca dari csv lain yang kemudian akan di predict oleh model yang telah di train karena masih kekurangan pengalaman dalam pembuatan.
2. Pengalaman pertama kali menggunakan library pada python dan juga belum terlalu biasa dengan syntaks python.
3. Banyak istilah dalam Neural Network yang belum kita ketahui sebelumnya.
4. Rata-rata pembahasan yang ada di Internet sulit untuk kami mengerti karena ilmu yang kami miliki masih cenderung dasar.

## **Solusi**

1. Untuk pembuatan model, solusi yang kami temui adalah dengan mencari melalui Google kemudian melakukan trial and error sehingga akhirnya
2. Mencari dokumentasi di internet dan juga banyak membaca serta belajar.
3. Mencari makna dari istilah tersebut di internet
4. Harus berusaha belajar lebih lagi dan merenungi apa yang telah dibaca.