LABORATORIUM PEMBELAJARAN ILMU KOMPUTER

FAKULTAS ILMU KOMPUTER

UNIVERSITAS BRAWIJAYA

BAB	: BACKPROPAGATION (1)
NAMA	: DIMAS TRI MUSTAKIM
NIM	: 205150200111049
TANGGAL	: 07/11/2022
ASISTEN	: ANDIKA IRZA PRADANA

A. Praktikum

1. Buka Google Colaboratory melalui tautan ini.
2. Tulis kode berikut ke dalam setiap *cell* pada *notebook* tersebut.
    a. Fungsi Binary Encoding dan Decoding

```
def bin_enc(lbl):
  mi = min(lbl)
  length = len(bin(max(lbl) - mi + 1)[2:])
  enc = []
  for i in lbl:
    b = bin(i - mi)[2:].zfill(length)
    enc.append([int(n) for n in b])
  return enc

def bin_dec(enc, mi=0):
  lbl = []
  for e in enc:
    rounded = [int(round(x)) for x in e]
    string = ''.join(str(x) for x in rounded)
    num = int(string, 2) + mi
    lbl.append(num)
  return lbl
```

    b. Percobaan Binary Encoding dan Decoding

```
labels = 1, 2, 3, 4
enc = bin_enc(labels)
dec = bin_dec(enc, min(labels))
print(enc)
print(dec)
```

    c. Fungsi One-hot Encoding dan Decoding

```
import numpy as np
def onehot_enc(lbl, min_val=0):
  mi = min(lbl)
  enc = np.full((len(lbl), max(lbl) - mi + 1), min_val, np.int8)
  for i, x in enumerate(lbl):
    enc[i, x - mi] = 1
```

```
  return enc

def onehot_dec(enc, mi=0):
  return [np.argmax(e) + mi for e in enc]
```

d. Percobaan Binary Encoding dan Decoding

```
labels = 1, 2, 3, 4
enc = onehot_enc(labels)
dec = onehot_dec(enc, min(labels))
print(enc)
print(dec)
```

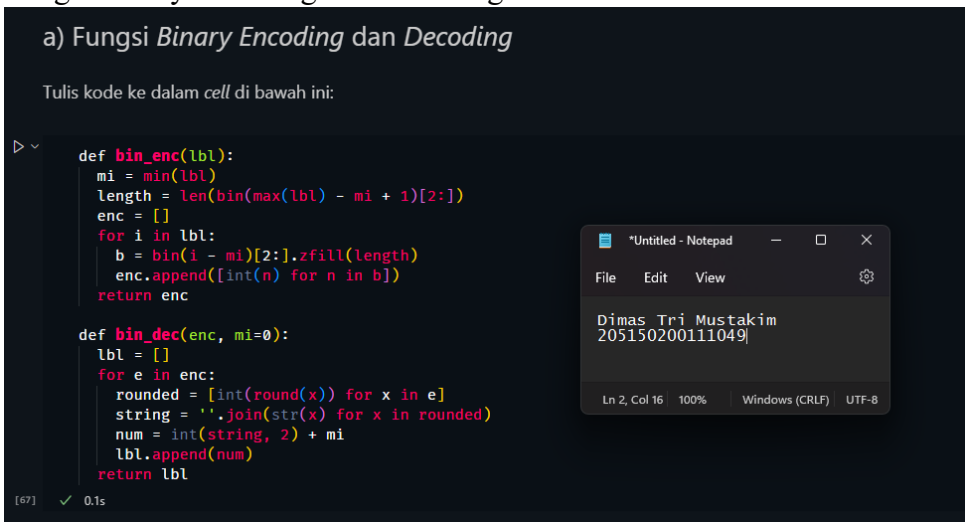
e. Fungsi Aktivasi Sigmoid dan Derivatifnya

```
def sig(X):
  return [1 / (1 + np.exp(-x)) for x in X]

def sigd(X):
  output = []
  for i, x in enumerate(X):
    s = sig([x])[0]
    output.append(s * (1 - s))
  return output
```

B. Screenshot

a. Fungsi Binary Encoding dan Decoding



b. Percobaan Binary Encoding dan Decoding

c. Fungsi One-hot Encoding dan Decoding

d. Percobaan Binary Encoding dan Decoding

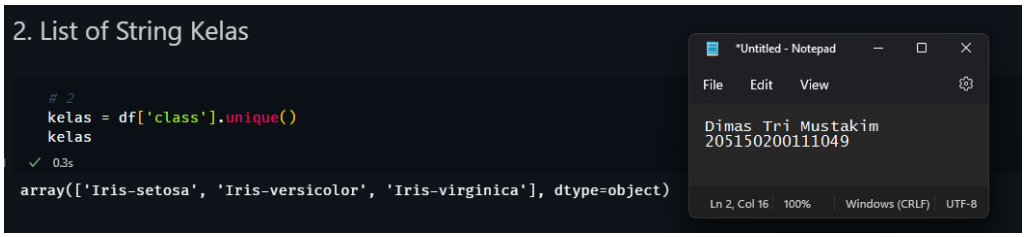
e. Fungsi Aktivasi Sigmoid dan Derivatifnya

C. Analisis

1. Download dataset Iris dalam format CSV di https://datahub.io/machine-learning/iris/r/iris.csv .
   Jawab:
   Untuk mengakses dataset Iris tersebut saya menggunakan fungsi read_csv dari library pandas.

2. Baca kolom terakhir pada file tersebut yang berisi kelas data. Buatlah variabel bernama kelas dengan tipe list of string. Variabel kelas berisi semua kelas yang terdapat pada file CSV tersebut.
   Jawab:

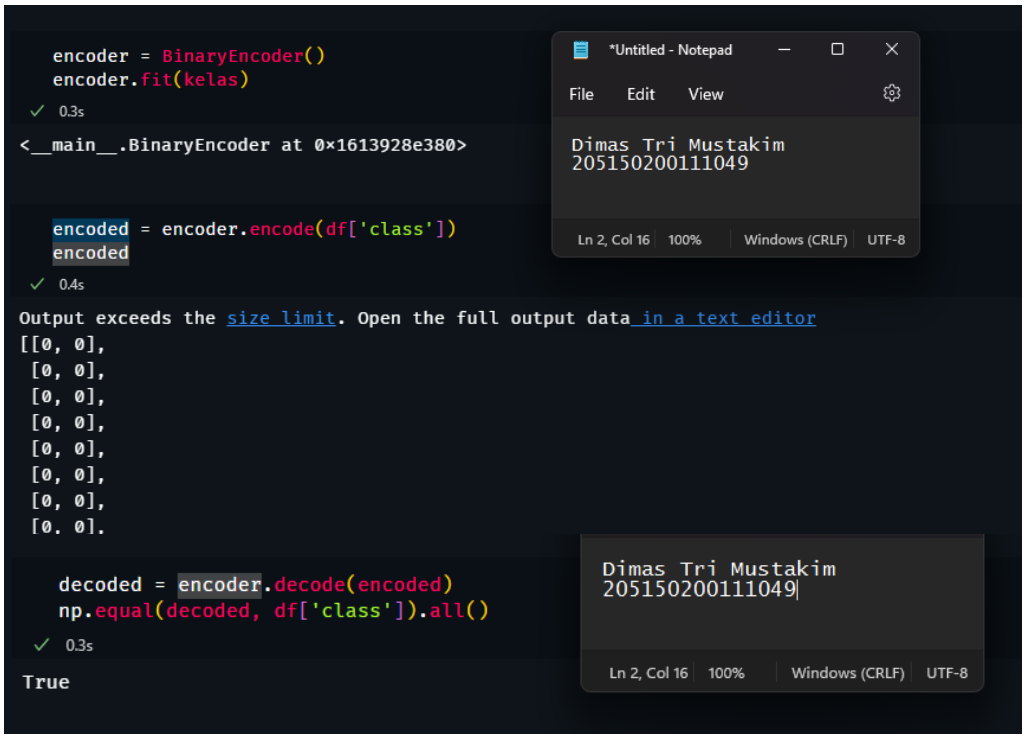
Untuk mengambil data pada kolom kelas dari pandas bisa menggunakan sintaks df[‘nama-kolom’], kemudian menggunakan fungsi unique untuk mengambil nilai yang unik saja.



3. Buatlah fungsi bernama bin_enc_str yang berfungsi untuk melakukan binary encoding pada string. Fungsi ini menerima input berupa list of string dan menghasilkan output berupa representasi binary encoding dari list tersebut. Jangan lupa membuat fungsi decodernya juga dengan nama bin_dec_str

Jawab:

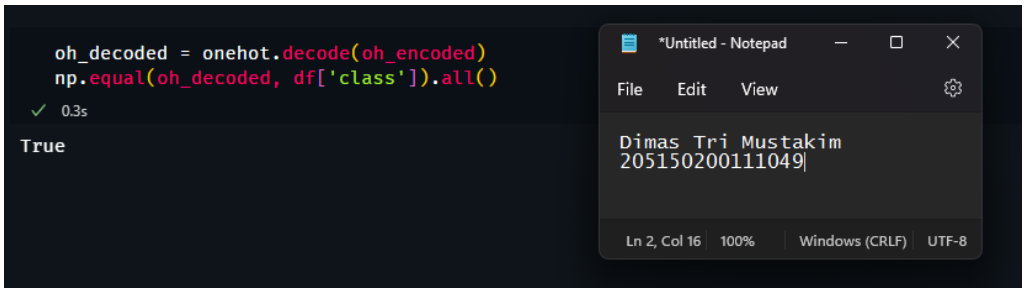
Fungsi tersebut saya buat di dalam kelas BinaryEncoder. Saya membuat kelas tersebut untuk kemudahan karena terdapat variabel yang perlu dibagi untuk fungsi encode dan decode agar bisa berfungsi dengan baik. Fungsi untuk encode saya beri nama encode, dan fungsi untuk decode sama namai decode. Saya rasa itu lebih bagus daripada bin_enc_str atau bin_dec_str. Data hasil encoding yang muncul untuk tiga kelas dalam data iris yaitu [0, 0], [0, 1], [1, 0].



4. Buatlah fungsi bernama onehot_enc_str yang berfungsi untuk melakukan one-hot encoding pada string. Fungsi ini menerima input berupa list of string dan menghasilkan output berupa representasi one-hot encoding dari list tersebut. Jangan lupa membuat fungsi decodernya juga dengan nama onehot_dec_str

Jawab:

Dengan alasan yang sama saya juga membuat fungsi untuk *one hot encoding* di dalam kelas dengan struktur fungsi yang sama. Data unik hasil encoding yaitu [1, 0, 0], [0, 1, 0], [0, 0, 1].



D. Kesimpulan

Backpropagation (backward propagation of errors) merupakan algoritma supervised learning untuk jaringan saraf tiruan yang menggunakan gradient descent. Algoritma ini akan menghitung error fungsi gradient sehubungan dengan bobot dari jaringan saraf tiruan. Algoritma ini dikenalkan pada tahun 1960-an dan baru 30 tahun setelahnya di populerkan oleh Rumelhart, Hinton dan Williams pada paper yang berjudul "Learning representations by back-propagating errors". Algoritma ini digunakan untuk melatih jaringan saraf tiruan melalui metode bernama chain rule. Secara sederhana, algoritma ini bekerja dengan cara melakukan backward pass selagi mengatur parameter dari model neural network setelah setiap feed forward yang melewati jaringan saraf tiruan. Algoritma backpropagation membutuhkan encoding karena arsitektur jaringan saraf tiruan ini memerlukan input dan output berupa angka. Hal tersebut dapat dilihat pada rumus yang digunakan seperti unutk pelatihan, tidak akan bisa dimasukkan string didalamnya.

Beberapa jenis encoding yaitu binary encoding, one-hot encoding, label/ordinal encoding, dan feature hashing. Binary encoding merupakan cara pengkodean dengan bilangan biner. Pertama-tama, kategori data diberi nomor urut dimulai dari nol. Kemudian, nilai binary encoding merupakan nilai biner dari nomor urut tersebut. One-hot encoding menggunakan sebuah vektor biner dengan panjang yang sama dengan banyaknya kelas atau kategori berbeda yang terdapat pada data. Salah satu bit pada

vektor biner akan bernilai 1 sedangkan lainnya 0. Posisi bit 1 menandakan kategori data. Ordinal encoding merupakan cara pengkodean dengan memberikan setiap kelas sebuah nilai numerik yang unik. Dan yang terakhir ada feature hashing yang melakukan pengkodean dengan cara menyandikan variabel kategori menggunakan bantuan fungsi hash dan disimpan dalam ruang dimensi tinggi menggunakan array ukuran tetap.