

1. Problem Statement:

FizzBuzz Task. In this task an integer divisible by 3 is printed as Fizz, and integer divisible by 5 is printed as Buzz. An integer divisible by both 3 and 5 is printed as FizzBuzz.

Solution

2) **Method 1 (Software1 method)** : Implemented the below fizzbuzz algorithm in python using if-else statements.

```
Result: fizzbuzz labelled integer
Initialization : low = 1, high = 100 ;
while  $low \leq high$  do
     $\left| \begin{array}{l} \text{if } low \% 15 = 0: \text{ print("fizzbuzz");} \\ \text{elif } low \% 3 = 0: \text{ print("fizz");} \\ \text{elif } low \% 5 = 0: \text{ print("buzz");} \\ \text{else: print(low);} \\ low - ; \end{array} \right.$ 
end
```

Algorithm 1: FizzBuzz Algorithm

3) **Method 2 (Software2 method)** : Use a neural network to learn fizzbuzz algorithm.

3.1) Data Pre-Processing

Training data : binary data from 101 to 1000.

Test data : binary data from 1 to 100.

3.2) Model Architecture

```
Model_1(
    (Layer_1): Linear(in_features=10, out_features=100, bias=True, activation = ReLU)
    (Classifier Layer): Linear(in_features=100, out_features=4, bias=True, activation = ReLU)
)
```

```
Model_2(
    (Layer_1): Linear(in_features=10, out_features=100, bias=True, activation = ReLU)
    (Layer_2): Linear(in_features=100, out_features=10, bias=True, activation = ReLU)
    (Classifier Layer): Linear(in_features=10, out_features=4, bias=True, activation = ReLU)
)
```

```

Model_3(
    (Layer_1): Linear(in_features=10, out_features=128, bias=True, activation = ReLU)
    (Layer_2): Linear(in_features=128, out_features=64, bias=True, activation = ReLU)
    (Layer_2): Linear(in_features=64, out_features=32, bias=True, activation = ReLU)
    (Layer_2): Linear(in_features=32, out_features=16, bias=True, activation = ReLU)
    (Classifier Layer): Linear(in_features=16, out_features=4, bias=True, activation = ReLU)
)

```

Loss Function : Cross-Entropy Loss

3.3) Model hyperparameters (1) :

| | |
|----------------|------|
| Optimizer | Adam |
| Learning Rate | 0.01 |
| Adam β_1 | 0.90 |
| Adam β_1 | 0.98 |
| Seed | 100 |

Result

| Model 1 result | | |
|----------------|-------------|-----------------------------|
| Batch Size | Epochs | Accuracy |
| 32 | 200 | 94.00 |
| 32 | 300 | 97.00 |
| 32 | 2000 | 98.00 (Best Model) |

| Model 2 result | | |
|----------------|--------|----------|
| Batch Size | Epochs | Accuracy |
| 32 | 200 | 90.00 |
| 32 | 500 | 95.00 |

| Model 3 result | | |
|----------------|--------|----------|
| Batch Size | Epochs | Accuracy |
| 32 | 200 | 79.00 |
| 32 | 500 | 81.00 |
| 32 | 1000 | 86.00 |

Model hyperparameter (2)

| | |
|--------------------|------|
| Optimizer | SGD |
| Learning Rate | 0.01 |
| Momentum β_1 | 0.90 |
| Seed | 42 |

Result

| Model 1 result | | |
|----------------|--------|----------|
| Batch Size | Epochs | Accuracy |
| 32 | 200 | 89.00 |

For other model types similar results is observed.

The accuracy of the best model is **98.00**. It was able to identify every fizz, fizzbuzz and buzz integer except 2 (integer 20 and 80).