Layers

```
class Layer():
    def __init__(self):
        self.params = []
        self.grads = []

    def __call__(self, x):
        return x

def backward(self, grad):
    return grad

def update(self, params):
    return
```

```python
class Linear(Layer):
    def __init__(self, d_in, d_out):
        self.w = np.random.normal(loc=0.0,
                                  scale=np.sqrt(2/(d_in+d_out)),
                                  size=(d_in, d_out))
        self.b = np.zeros(d_out)

    def __call__(self, x):
        self.x = x
        self.params = [self.w, self.b]
        return np.dot(x, self.w) + self.b

    def backward(self, grad_output):
        grad = np.dot(grad_output, self.w.T)
        self.grad_w = np.dot(self.x.T, grad_output)
        self.grad_b = grad_output.mean(axis=0)*self.x.shape[0]
        self.grads = [self.grad_w, self.grad_b]
        return grad

    def update(self, params):
        self.w = params[0]
        self.b = params[1]
```