

input_layer

```
graph TD; A[input_layer] --> B[Flatten]; B --> C[Dense  
kernel <784x128>  
bias <128>  
Activation]; C --> D[Dense  
kernel <128x10>  
bias <10>]; D --> E[dense_1];
```

Flatten

Dense

kernel $\langle 784 \times 128 \rangle$

bias $\langle 128 \rangle$

Activation

Dense

kernel $\langle 128 \times 10 \rangle$

bias $\langle 10 \rangle$

dense_1