MNIST Adam

model

```
model = Sequential([
        tf.keras.layers.Dense(units=25, activation='sigmoid')
        tf.keras.layers.Dense(units=15, activation='sigmoid')
        tf.keras.layers.Dense(units=10, activation='linear')
])
```

compile

$\alpha = 10^{-3} = 0.001$

```
model.compile(optimizer=tf.keras.optimizers.Adam(learning_rate=1e-3),
    loss=tf.keras.losses.SparseCategoricalCrossentropy(from_logits=True))
```

fit

```
model.fit(X,Y,epochs=100)
```

The Adam optimizer is the recommended optimizer for finding the optimal parameters of the model. How do you use the Adam optimizer in TensorFlow?

- ● When calling model.compile, set optimizer=tf.keras.optimizers.Adam(learning_rate=1e-3).

- ○ The Adam optimizer works only with Softmax outputs. So if a neural network has a Softmax output layer, TensorFlow will automatically pick the Adam optimizer.

- ○ The call to model.compile() uses the Adam optimizer by default

- ○ The call to model.compile() will automatically pick the best optimizer, whether it is gradient descent, Adam or something else. So there's no need to pick an optimizer manually.

✓ **Correct**

Convolutional Neural Network
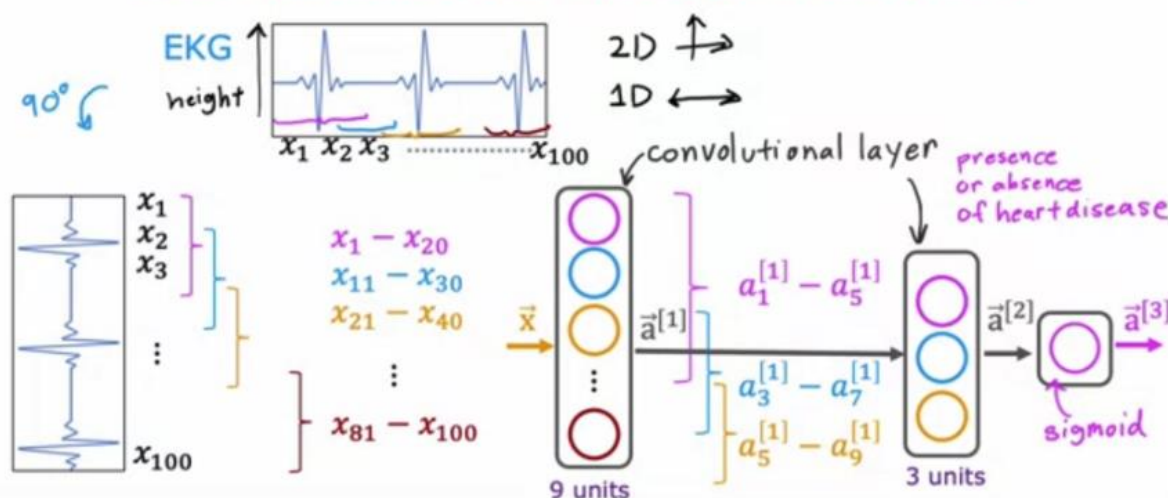
The lecture covered a different layer type where each single neuron of the layer does not look at all the values of the input vector that is fed into that layer. What is this name of the layer type discussed in lecture?

- ○ 1D layer or 2D layer (depending on the input dimension)
- ● convolutional layer
- ○ A fully connected layer
- ○ Image layer

✓ **Correct**
Correct. For a convolutional layer, each neuron takes as input a subset of the vector that is fed into that layer.