**Section 1**

**Aim:**

Prediction of male and female names by creating a neural network model using Long Short-Term Memory and Convolution Neural Network. And then combine these two models to get a new model. In the end, we have to compare the accuracies of these models.

**Introduction:**

We will have two files as part of the nltk names data and they are male.txt file which contains male names and female.txt file which contains female names. The names are then shuffled at random, and the data is divided into train and test data. The data must then be preprocessed. The data must then be vectorized. In Task 1, we have to build a convolutional neural network model consisting of two 1D convolutional layers. After creating the convolution layers, we add a sigmoid activation with a dropout of our desired unit. Finally, we add a softmax layer. In Task 2, we must create a sequential model with a 32-unit LSTM layer. Coming to Task 3, we combine task 1 and task 2 to build a model.

**Section-2**

**Explanation of LSTM and CNN:**

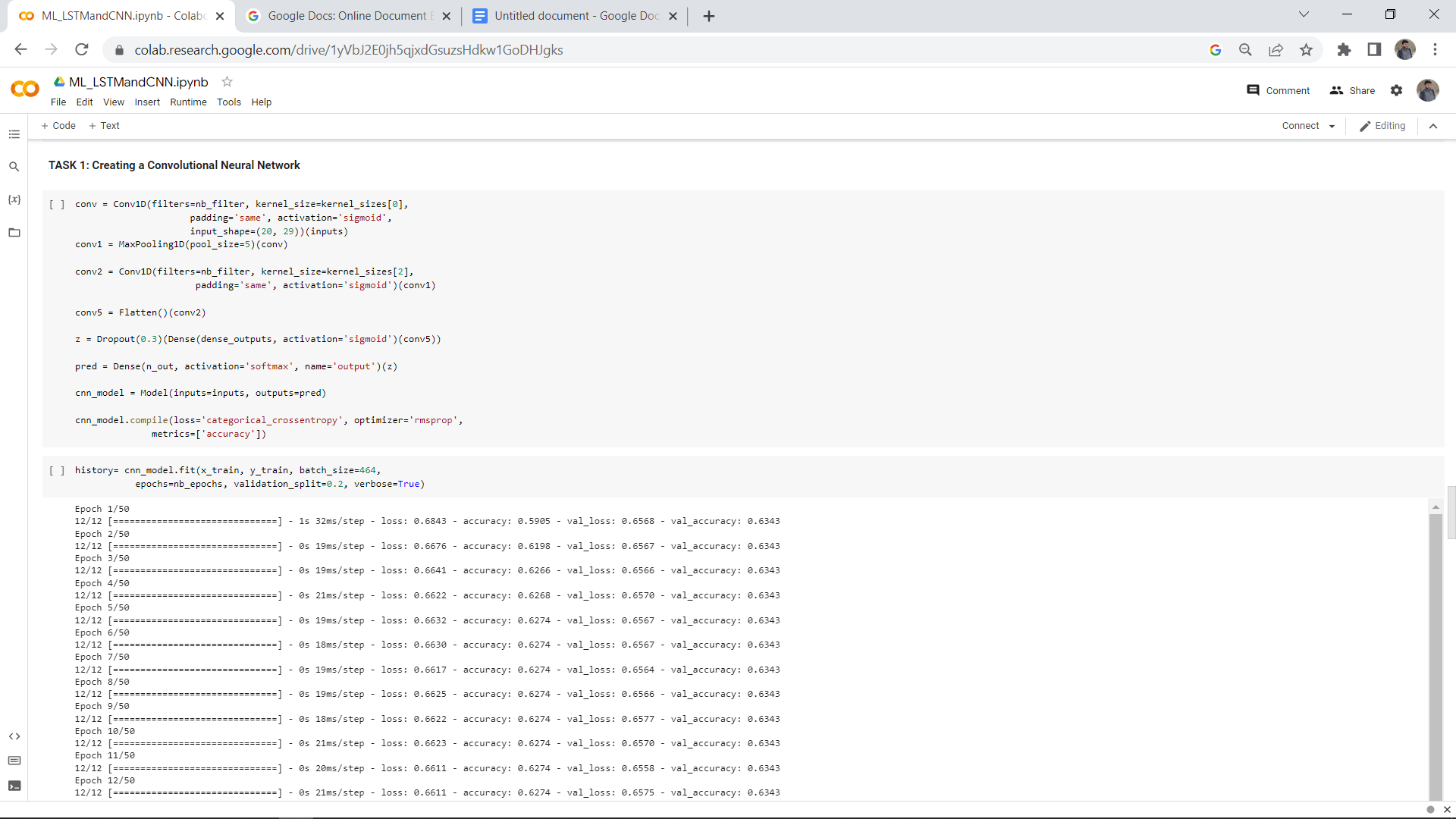
Long Short-term Memory Network:LSTMs employ a series of 'gates' that regulate how information in a data sequence enters, is stored in, and exits the network. A typical LSTM consists of three gates: a forget gate, an input gate, and an output gate. These gates are like filters, and they each possess their own neural network. Long Short-Term Memory (LSTM) networks are a class of recurrent neural networks that can learn order dependence in sequence prediction tasks. This is a necessary behavior in complex problem domains such as machine translation, speech recognition, and others. LSTMs are a complicated branch of deep learning. It can be difficult to grasp the concept of LSTMs and how terms like bidirectional and sequence-to-sequence relate to the field. The cell state and its various gates are at the heart of LSTMs. The cell state serves as a transportation highway for relevant data all the way down the sequential chain.You can believe of it as the network's "memory." In theory, the cell state can carry relevant data throughout the sequence's processing. As a result, information from earlier time steps can travel to later time steps, diminishing the effects of short-term memory. As the cell state travels, information is added or removed from the cell state via gates. The gates are various neural networks that determine which information on the cell state is permitted. During training, the gates can learn what information is important to keep or forget.

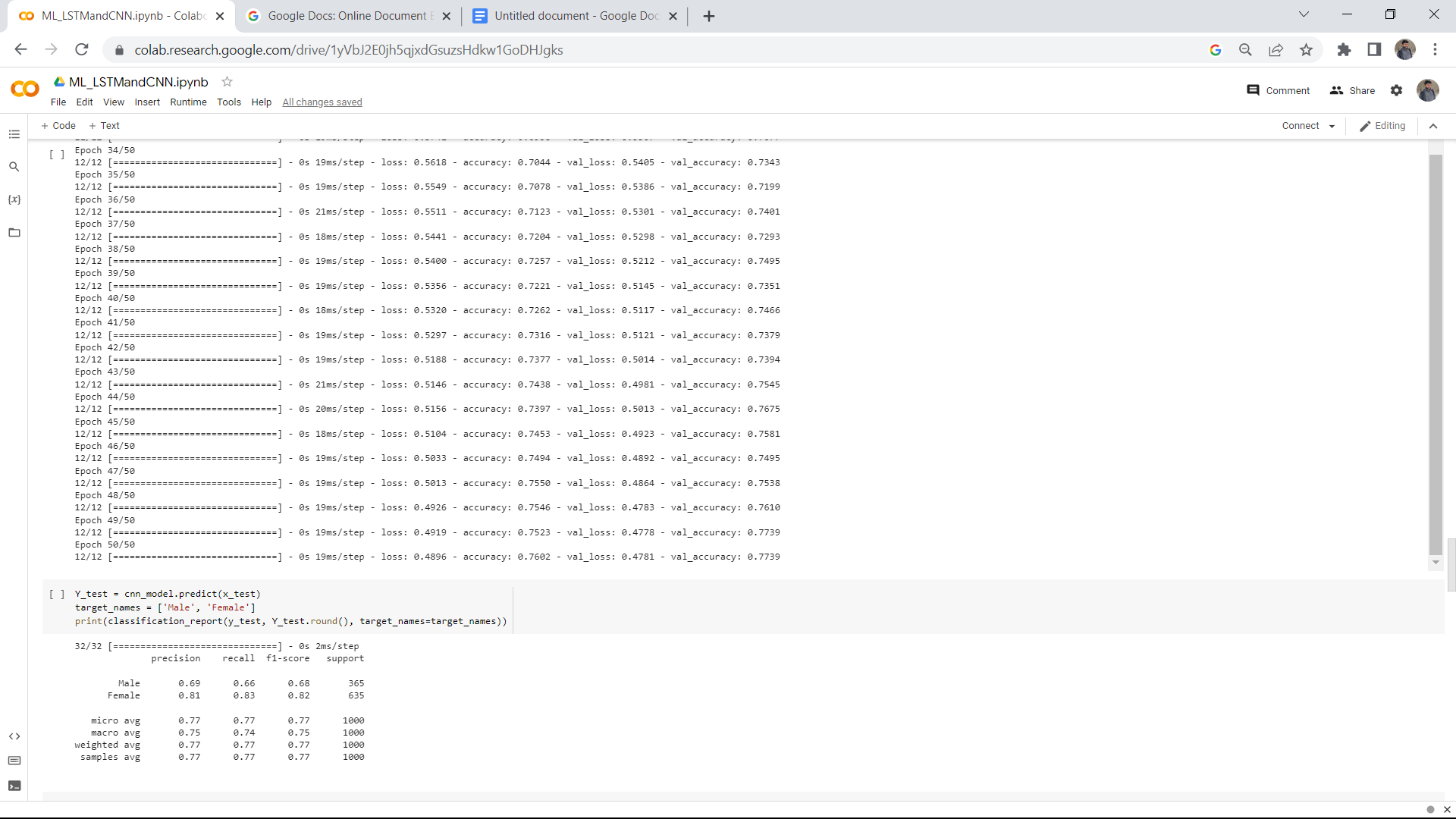
Convolutional Neural Networks (CNN): CNNs were created and put to use for the first time in the 1980s. At the time, a CNN could only recognize handwritten numbers to a certain extent. To read zip codes, pin numbers, etc., it was mostly utilized in the postal industry. The most crucial thing to keep in mind about any deep learning model is that it needs a lot of computational power and data to train. Because of this significant disadvantage at the time, CNNs were restricted to the postal industry and were unable to enter the machine learning field. A Convolutional Neural Network (ConvNet/CNN) is a Deep Learning method that can take in an input image, give various elements and objects in the image importance (learnable weights and biases), and be able to distinguish between them. Comparatively speaking, a ConvNet requires substantially less pre-processing than other classification techniques. ConvNets have the capacity to learn these filters and properties, whereas in primitive techniques filters are hand-engineered.

**Section-3**

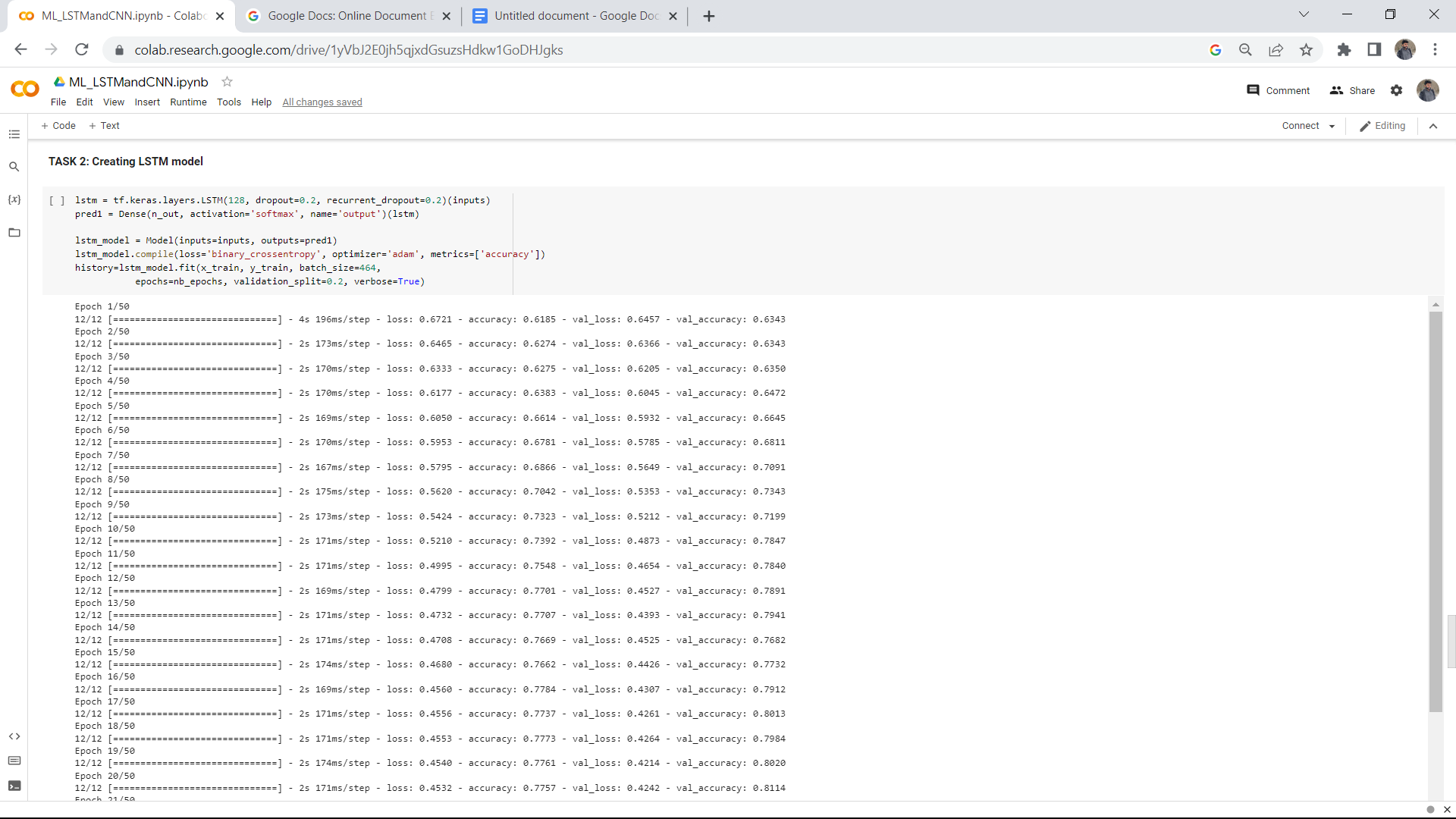
**Explanation of all tasks:**

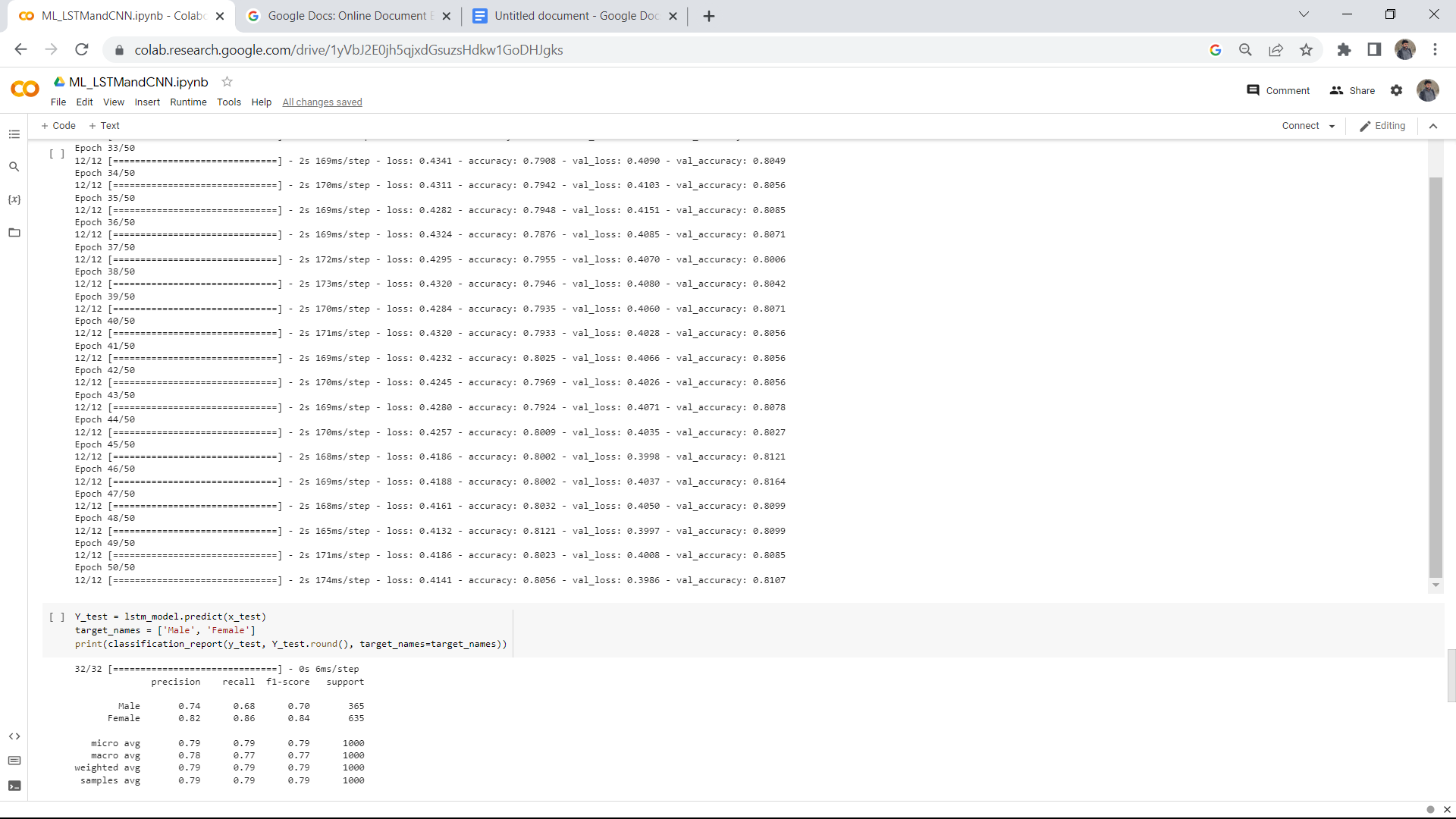
We are going to build a convolutional neural network model for **Task 1** that has two 1D convolutional layers. After generating the convolution layers, we add a sigmoid activation by using a dropout of 0.3. Finally, we build a softmax layer.



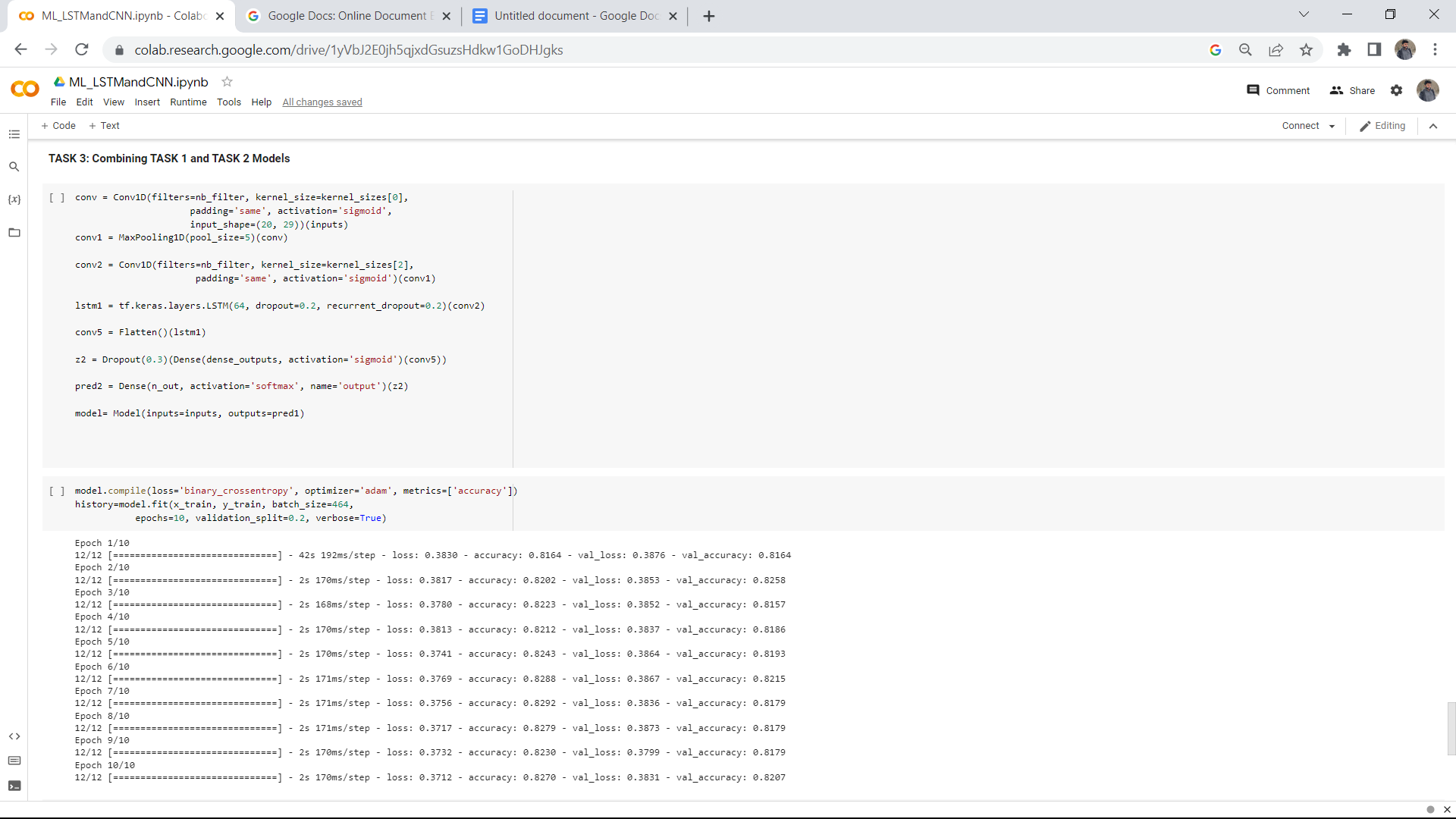


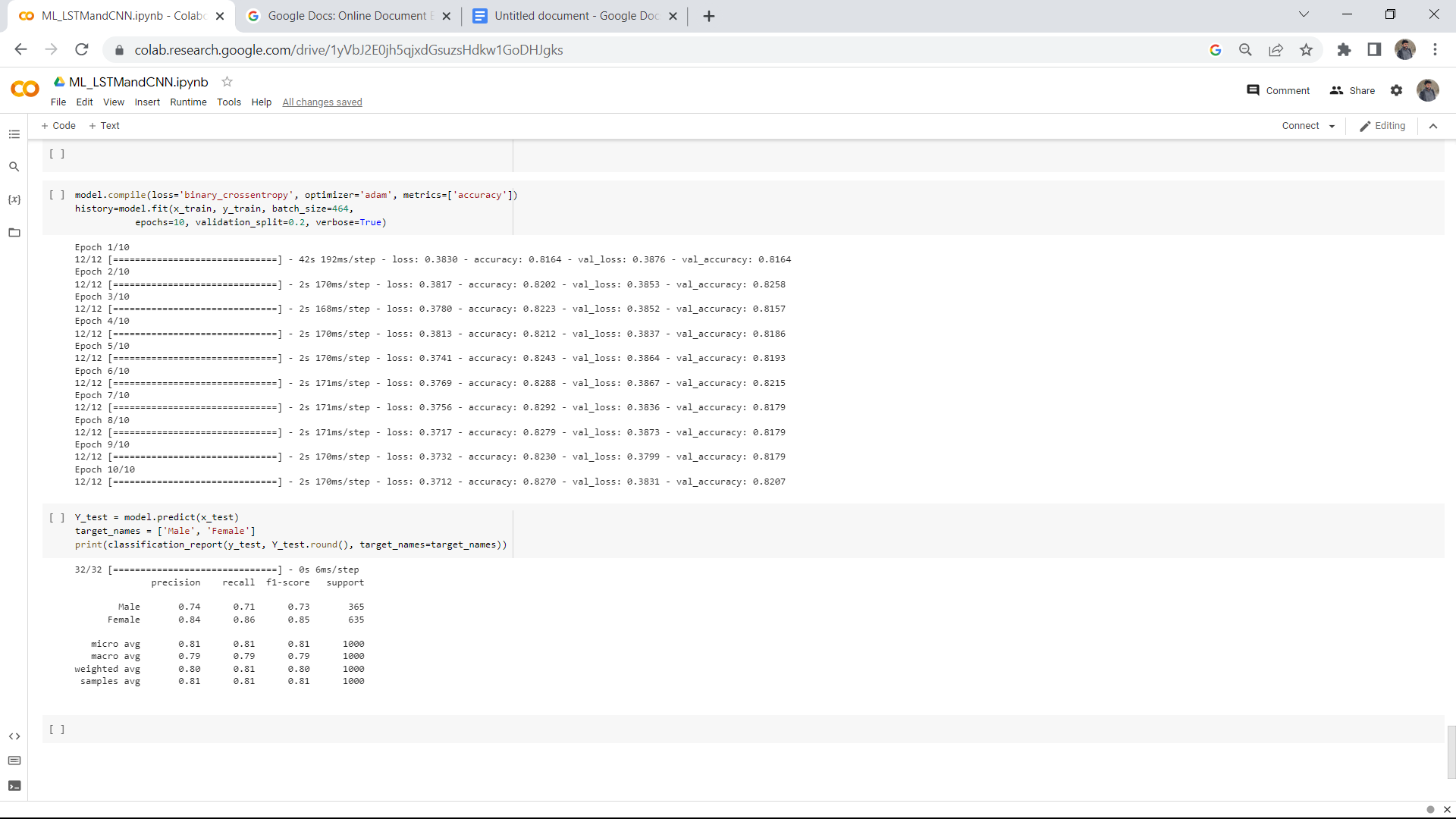
In **Task 2,** we are going to create a model using Long Short-term Memory





Task 3, combining both LSTM and CNN,





**Section-4**

**Comparing the Models:**

* In task-1, after 50 iterations the model’s accuracy is approximately 77%.
* When compared to task-1, Task-2’s model accuracy increased to 80%.
* Now compared to task-1 and task-2, in task-3, the model attained an accuracy of 80% just for 10 iterations.

**Section-5**

**Conclusion:**

As opposed to a CNN, which is made to take advantage of "spatial correlation" in data and performs well on images and voice, an LSTM is typically used to process and generate predictions given sequences of data. An LSTM is essentially a neuron unit that feeds data back to itself for the following time step in a series. It is a variant of recurrent neural networks, a more diverse family of neural networks.