

**Adam optimizer**

**Learning rate: 0.03**

**Vocabulary size – 2000**

**Num of steps 20 (Also tried with no of steps as 1000)**

**The nodes are scattered very long**

**The higher the learning rate, the faster the model trains the data**

**Tensor board**

