# 从Keras到浏览器的情感分类

我们将使用Keras创建一个分类器来区分电影中的积极情绪和消极情绪。然后我们将模型加载到浏览器中进行用户预测。

#### 守则

我在上创建了一个存储库，其中包含遵循本教程所需的代码。如果您发现任何错误或遇到任何问题，请在此提出问题。

#### 数据集

我们将使用进行情绪分类。数据集包含7086条关于带标签电影的语句。标签1表示积极情绪，0表示消极情绪。

#### 数据预处理

对于文本分析，我们首先需要对数据进行预处理。首先，在上传数据时，我们需要将句子拆分成单词，然后填充训练集。

然后，我们需要创建字典单词索引，将单词映射到整数以便嵌入。注意字典的长度可能很大，所以我们计算训练集中每个单词的计数，并且只包括重复次数超过某个阈值的单词。

process方法将通过删除标点符号从语句中创建单词

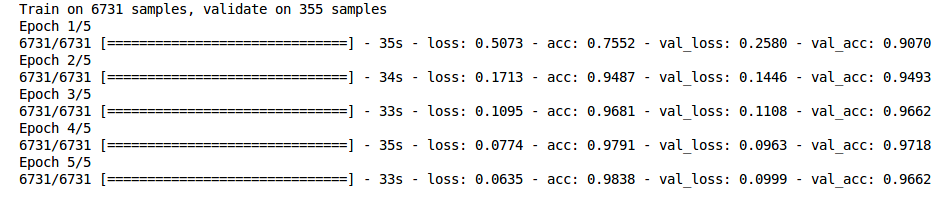
之后，我们需要将每个感伤语句映射为一个整数序列，但首先要得到包含最大单词数的语句

现在，我们可以使用字典word\_index创建一个整数序列注意，我们创建一个长度为max\_tokens的固定数量的序列，该序列具有最大数量的token。我们使用预填充，因为它可以提供更好的结果

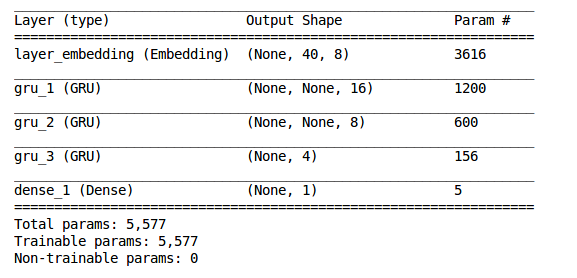
现在我们准备创建keras模型。第一层是嵌入层。然后我们用3层来跟踪它。最后我们创造了一个致密的乙状结肠激活层。我们使用Adam优化器编译模型

然后我们对模型进行5个阶段的训练，5%的分割用于32个批量的验证

验证集的准确率达到97%。



我们可以使用



我们将使用.h5格式保存模型

现在我们已经完成了模型，我们将导入它并在浏览器上运行。首先，我们需要将其转换为json格式。在此步骤之前，您需要使用

将keras模型转换为tensorflowjs可以理解的模型

这将创建一个json文件，其中包含元变量和其他一些名为group1-shard1of的变量，其中包含权重的计算值

#### 将模型移植到浏览器

为了便于处理，我们将加载由keras代码生成的字典

因此，word\_index现在将包含与上一节中相同的单词、索引对。之后，我们需要创建一些helper方法来处理文本，将其标记化，然后使用字典将其映射到整数

现在我们可以使用tensorflow.js将其加载到浏览器中

Finally we combine all the steps using one method

Using the sample text we get a prediction value 0.98which is very close to 1.

#### References

[**Hvass-Labs/TensorFlow-Tutorials** TensorFlow-Tutorials - TensorFlow Tutorials with YouTube Videosgithub.com](https://github.com/Hvass-Labs/TensorFlow-Tutorials)

[**tensorflow/tfjs** tfjs - A WebGL accelerated, browser based JavaScript library for training and deploying ML models.github.com](https://github.com/tensorflow/tfjs)