

1. Overview

After a few more meetings, your team has been assigned to address the following issues asked by the stakeholders:



ZHAO, CEO OF WELCOMEBIKE

“ So, there's not a lot to discuss at this point. We need a neural network that can predict the number of bikes we can expect to rent on a given day. I'll leave it up to you to decide which factors matter there.



JOHNNY, THE DATA SCIENCE INTERN

“ Given how many variables there are in building neural networks, and how long it takes to train and test them, we probably need everyone on the team working on building and testing neural networks. ”



I'd suggest that you coordinate how to tackle that in a smart way.



ZHAO, CEO OF WELCOMEBIKE

“ Oh one more thing, while you've been working with data up through October for your analysis, we just received November and December's as well, and I'd like to see how your model does with it. ”

Could you give me the total number of bikes you'd expect for each day in this dataset? I'm not concerned with whether they are registered or casual at this point, just the total number for each hour of each day found here:

[bikes_december.csv](#).

Please make sure to submit both a .csv with these predictions and your executive summary.

2. More Tips from Johnny

⚠ KERAS VS TF.KERAS

As [mentioned in the reading](#), Keras used to be a standalone library, but as of September 2019, it is part of Google's TensorFlow library.



Keep that in mind if you're looking at any tutorial that was written prior to that date. Most of the API and functions will be the same, but your import statements will likely be different.

For more information, [see this article on the change](#).

i DATA DICTIONARY

Use this [data dictionary](#) to help explain the values and sources of different columns in the dataset.

i FEATURE SCALING

If you're going to be comparing different numeric features, be sure they are using the same scale. You may find it useful to use min-max scaling to handle this problem. You could do this calculation manually, or use [Sci-Kit Learn's MinMaxScaler](#)

i KERAS & SCI-KIT LEARN

Just because you're using keras and tensor flow to build the actual neural network, that doesn't mean you can't take advantage of some of the preprocessing and analysis modules in sci-kit learn.

Keras also provides [two wrappers](#) that allow you to use sklearn pipeline style code with keras neural networks.

For more information, see:

- [How to insert Keras model into scikit-learn pipeline?](#)
- [How to Grid Search Hyperparameters for Deep Learning Models in Python With Keras](#)

i TRAINING CHECKPOINTS

Neural networks can sometimes take a long time to train (especially once we move into using image data in future modules). If you find you can't let the model train for as long as is needed in a single run, you might want to look into [training checkpoints](#).

This [tutorial](#) may be easier to follow, but remember to update the import statements as needed.



i GRID SEARCH

There are a lot of hyperparameters to test out. You may want to use either a grid search or a random grid search. Here is an example in Google Colab.



JOHNNY, THE DATA SCIENCE INTERN, CATCHES YOU AFTER WORK:

“ Hey, I know you're probably busy, I was going to put together some hints and sample code for you like I usually do, but the best example code I can provide is what you see in those [keras tutorials from the reading assignment](#). So, be sure to read over those.

-
1. [CEO photo by Sung Wang on Unsplash](#) ↵
 2. [Investment Banker photo by steffen Wienberg on Unsplash](#) ↵
 3. [Data Science Intern photo by Fábio Lucas on Unsplash](#) ↵