**SIESTA Protocol**

Thanks for trying out SIESTA for us! The protocol below should have all the information you need to get started. If you have any questions or doubts, email Ray ([raysan53@uw.edu](mailto:raysan53@uw.edu)) and Carlos ([caldart@uw.edu](mailto:caldart@uw.edu)) and we will get back to you as quickly as possible.

**Downloading Python**

If you don’t already have Python installed on your machine, start here. If you already have Python, proceed to Step 3.

1. Download Python 3.7.7 for your operating system at this link, and follow the instructions for installation: <https://www.python.org/downloads/>
2. When installation is complete, you should have a folder called “Python 3.7” on your machine.

**Setting up Python to Run SIESTA**

SIESTA makes use of several open-source Python packages for analysis. We have tried to set up the download folder so that all of these packages should be included, but Python can be finnicky, so the following instructions are designed to make sure you have access to all of the dependencies necessary to run the program. This will not be an issue for the final version, which will simply be an executable application that you can click and run from your Desktop.

1. Make sure you have downloaded the folder we have provided you titled “Siesta V5”. It should contain everything you need to get started.
2. Open a Terminal window on your machine to install the following Python packages on your computer. The commands are slightly different depending on if you are using a Mac or Windows machine. Type the following lines of code and press enter, one at a time:
   1. **Scikit-learn, numpy and scipy.** These are the packages used to perform machine learning analyses.

*Windows: Enter the below commands into your Terminal:*

pip install -U scikit-learn

pip install -U scipy

pip install -U numpy

*Mac - Enter the below command into your Terminal:*

pip install -U numpy scipy scikit-learn

* 1. **PyEDFlib**. This package is used to read in raw EEG/EMG data from .edf files into Python

*Windows and Mac – Enter the below command into your Terminal:*

pip install pyEDFlib

* 1. **Kivy**. This package allows Python to render the graphical user interface (GUI) that you will use to score sleep data. Installation of kivy can be a little bit tricky – follow the next steps very carefully.

*Windows* – Go to the following link and follow all of the instructions in the “Installing the kivy stable release” section. <https://kivy.org/doc/stable/installation/installation-windows.html>

*Mac**- Enter the below command into your Terminal:*

Python -m pip install kivy

* 1. **Pandas**. Pandas is a package for data manipulation in Python.

*Windows and Mac – Enter the below command into your Terminal:*

pip install pandas

1. Make sure that ALL of the above packages installed successfully – without them, you will not be able to run SIESTA. If you have any issues with installation or doubts, get in touch with us before you run the program. Otherwise, proceed to Step 6.

**Running SIESTA with the de la Iglesia Lab Sleep Database**

You are now ready to run SIESTA. Before you proceed, double check that you have downloaded all the contents of the “Siesta V5” folder and that the data you will use for scoring is in EDF file format. The following set of instructions will walk you through loading in your raw EEG/EMG data, extracting signal features from it, and automatically scoring it using the database that we have provided.

1. Navigate to your “Python 3.7” folder open the application “IDLE”.
2. With IDLE open, click “File” -> “Open” and navigate to the file in the “Siesta V5” folder we have provided, and select the file “main.py”. Your screen should now look something like this:

**A screenshot of a cell phone

Description automatically generated**

1. In the top navigation window of IDLE, select “Run” -> “Run Module”. If successful, you will now see the SIESTA GUI appear, as below:

**A screenshot of a cell phone

Description automatically generated**

1. Adjust the settings in the **“SiestaConfig”** window to meet the specifications of your .edf files. By default, the de la Iglesia lab samples at 400 Hz, reads in EEG data from Channel 2 of the EDF file, reads EMG data from Channel 1 of the EDF file, and scores data in epochs of 10 seconds long. Also note that SIESTA makes use of signal features in the 0.5 Hz to 50 Hz range – if you do any filtering to your data that cuts of frequencies outside this range, this may affect SIESTA’s performance.
2. Click the right arrow icon in the upper righthand corner of the SIESTA window. You should now see a module called **SiestaFeat,** that looks like the window below. This module is the first step of SIESTA, where the algorithm will calculate features of the raw EEG and EMG data that will later be used in scoring.

A screenshot of a cell phone

Description automatically generated

1. Select “Load” and navigate to the EDF file that you want to score. This will trigger SIESTA to calculate all the features necessary for scoring in epoch lengths you specified in SiestaConfig. Depending on the length of your recording, this may take a couple of minutes. Once the file is loaded, click “Save” and name the features file whatever you wish. SIESTA will create a CSV file with all of these features.
2. Before proceeding, check that the features file has been created. You can open it up in Excel or any text editing program.
3. Click the right arrow again to proceed to the **SiestaScore** module, skipping over **SiestaFit** for now. We will come back to that module later. **SiestaScore** allows you to load the training dataset that will be used for automatic sleep scoring.
4. Click “Load Database” and navigate to the “Siesta” subfolder in the “Siesta V5” folder we have provided. Select the file called “Siesta\_Model.sav”. You should see the field below populate with information about the database you have just loaded.
5. Next, click “Load Mice” and select the feature file that you generated in Steps 11 and 12. When the field below populates with an image of your raw EEG data, the file has loaded successfully.
6. Finally, click “Save” and name the final scored output file whatever you wish. If you completed Steps 13-16 successfully, your screen will look something like this:

**A screenshot of a cell phone

Description automatically generated**

1. Check your final scored output file in a spreadsheet viewer or text editor. You should have timestamped scores (either AWAKE, NREM or REM) for each epoch of your recording, along with values for commonly used in analyses of sleep including delta power, theta power and sleep spindle frequency band power.
2. ***You are done!*** Use this output file for your analyses or score more data. This is the general process for scoring with SIESTA. Read the steps below for a couple of different scenarios that may also be useful depending on your needs.

**Running SIESTA with Your Own Manually Scored Data**

While we have tried to provide a large and diverse set of training data for scoring with SIESTA, we also give users the option to load their own data. If you have a large set of manually scored data from your own lab that you want to use to score data, follow the steps below.

1. Extract signal features from your raw data using **SiestaFeat**, as described in Steps 10-12 in the previous section. This should result in the creation of a new Feature file. Do this for as many files as you wish to include in your database.
2. Create a separate CSV file containing the scores for your manually labeled data. Save the Feature file and Scores file in the same folder. Do this for as many files as you wish to include in your database.
3. Navigate to the **SiestaUser** window in the SIESTA GUI. Select “Load” and select the folder that contains the feature and score file.
4. Next, click “Save” and name the file whatever you wish. You will now have a CSV file combining all features and scores into a single training dataset.
5. Proceed to the **SiestaFit** module. Click “Load” and select the CSV file containing the training dataset you have just generated.
6. SIESTA will now train on your data. Depending on the size of your dataset, this may take a few minutes. Once this is complete, you will see the left field under “Load” populate with information about the model.
7. Click “Save” and save the model with whatever name you wish, followed by the file extension “.sav”. For example, “My\_Database.sav”. This stores the data in a format that does not require you to re-train SIESTA every time you want to score new data.
8. Proceed to Step 13 in the previous section and follow the instructions to finish scoring your data, loading your new .sav file instead of the “Siesta\_Model.sav” file that we provide.

**Running SIESTA with de la Iglesia Lab Sleep Database AND Your Own Data**

While SIESTA classifies new raw data with an average scoring accuracy of 88% between all 3 stages using our database alone, our characterizations of SIESTA found that performance increases substantially if the user adds a small bit of manually scored data from the target file to the database. For example, if you have a 24-hour recording to score, and you manually score 2 hours of that recording and add it to our training dataset, your performance accuracy could increase by as much as 3%. Users can also add their own manually scored data to the training dataset in this fashion.

1. Create a new CSV file containing your training data using **SiestaUser** by following Steps 19-22 in the previous section.
2. Using a spreadsheet program or text editor, combine your dataset CSV file with the CSV file we provide titled “Training Dataset.csv”. Save your new file with whatever name you choose.
3. Proceed to step 23 of the previous section and follow the instructions to continue scoring your data.