6002B Machine Learning CW Setup:
Getting started with `tsml` in GitHub Desktop and IntelliJ

This guide is written for users to download the **tsml** code base and switch to the **ml6002b-coursework** branch using GitHub Desktop. It is intended for users to follow if they want to make sure they are downloading the code correctly or are new to git/GitHub. If you are already familiar with git/GitHub, then feel free to follow your usual workflow to clone tsml and switch to the ml6002-coursework branch.

Pre-requisites
You need to:

1. have a **GitHub account.** You can use an existing account if you have one or you can sign up at https://github.com/join
2. download **GitHub Desktop**. This is a desktop client that makes cloning repositories (and changing branches within them) easy. It is a free download from https://desktop.github.com/

Once you have installed GitHub Desktop you will need to make sure that you are **logged in with your GitHub account** and that **Git is correctly configured**.

Sign in: If you are not prompted to sign in when you first open GitHub Desktop, you can do this by pressing **File -> Options -> Accounts**.
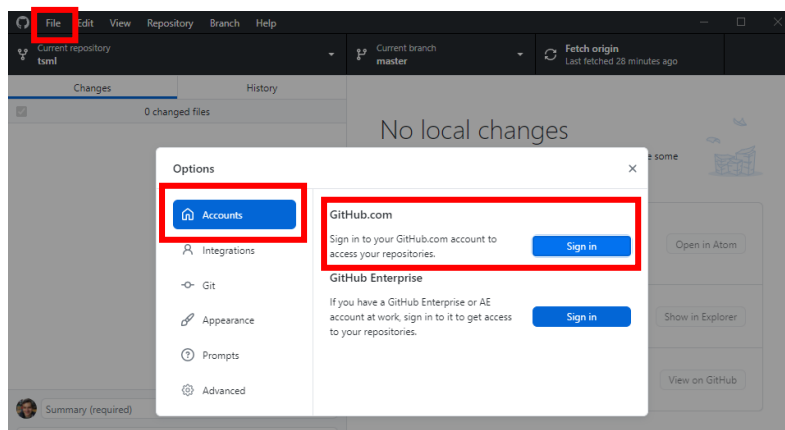
Configure git: If you are not prompted to configure git when you first open GitHub Desktop, you can do this by pressing **File -> Options -> Git**. The important thing is that the email address matches your login:
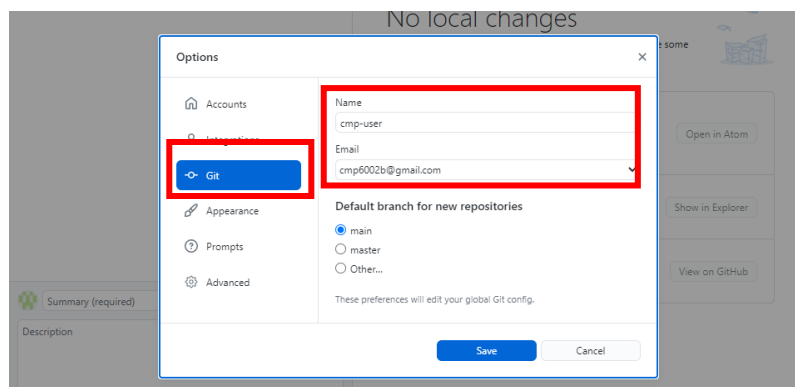
Cloning `tsml` and Switching Branches

You will be using a toolkit called tsml (https://github.com/uea-machine-learning/tsml), which is a Weka-compatible Java toolbox for time series classification, clustering and transformation that has been/is being developed by Tony, Jason, and the rest of the time series research group at UEA.

For the coursework assignment you will need to **clone the repository** and **switch to a specific branch** that has been created for the coursework assignment
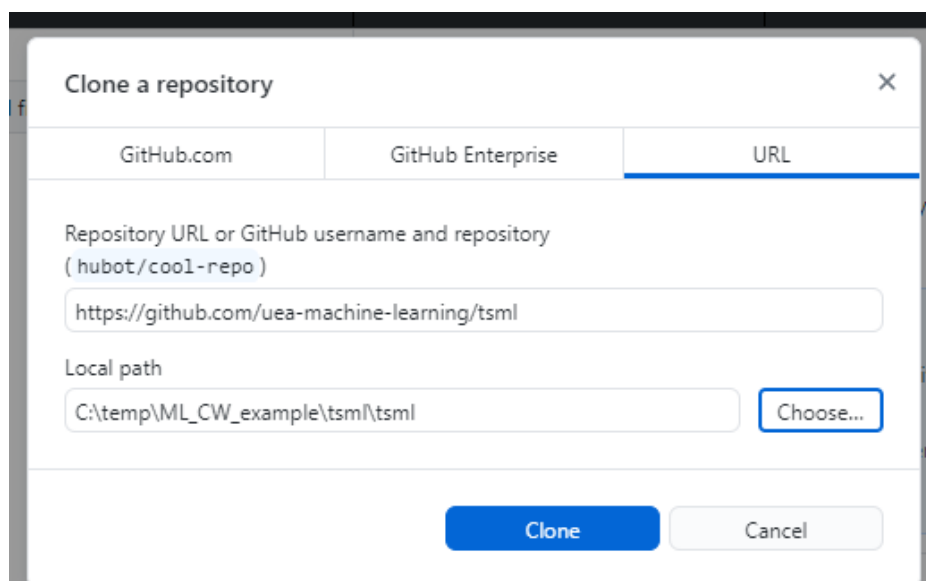
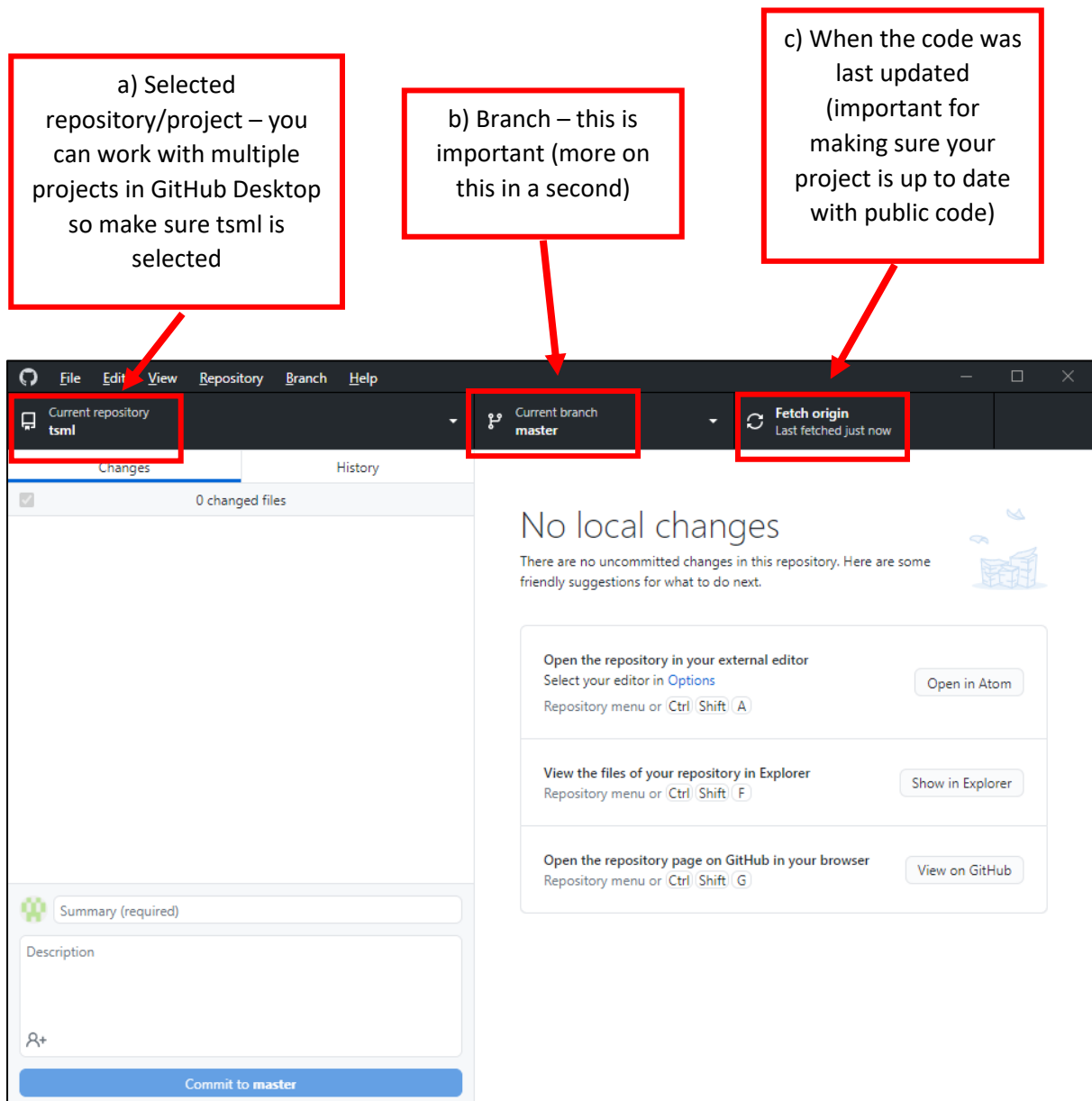## Cloning `tsml`

(*instructions modified from docs.github.com*)

1. In the **File** menu, click **Clone Repository**
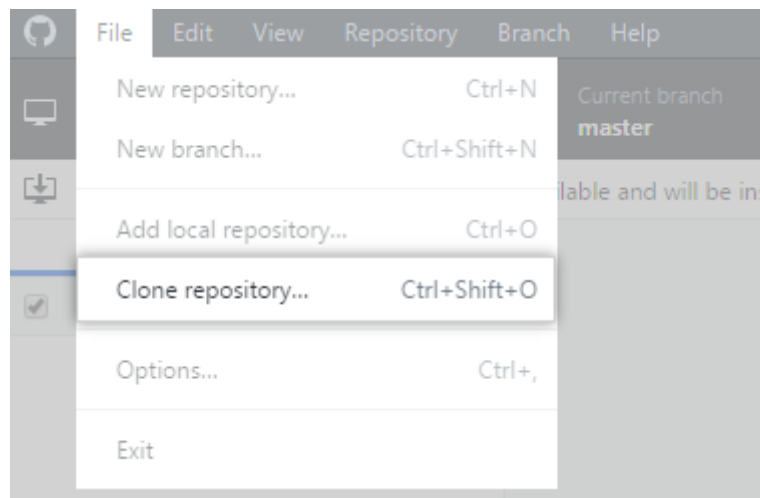
2. Click the **URL** tab and enter the URL for tsml: https://github.com/uea-machine-learning/tsml. Set the local path to where you would like the code to be downloaded to on your machine. Click clone and the download will start.

3. Once it has finished cloning you will see the following:

a) Selected repository/project – you can work with multiple projects in GitHub Desktop so make sure tsml is selected

b) Branch – this is important (more on this in a second)

c) When the code was last updated (important for making sure your project is up to date with public code)

Switching to the `ml6002b-coursework` branch

GitHub repositories can have multiple **branches**. The idea behind branches is that different features can be developed simultaneously for the same code base without changes clashing with one another. The simplest projects having a single branch called master (this is being phased out and renamed as main in new repositories, however). A branch can be made that uses main/master, allowing a user to change the code in the branch without affecting the base. Once all changes are completed, the new branch can be merged back into main/master to update it with the new features, and then the separate branch is closed.

You do not need to know too much about branches for the coursework assignment except that there is a specific branch that has been made for the assignment. You will need to change the **Current branch** selection in the screenshot above (highlighted as b) **from master** to **ml6002b-coursework**. This is a

branch that has been prepared for the assignment. Switching to this branch will update the code that is saved on your machine to match what is in the ml6002b branch,

It is also good practice to *pull* code regularly to make sure that the base code is up to date – this can be done by pressing the refresh button next to **Fetch origin** that is highlight as c in the screenshot on the previous page. Please note that you will never need to push anything to the tsml project/ml6002b-coursework branch – your work should be done in private and you should just use the code from this branch as a base (you may be asked to pull/update the origin however if any changes are made to the underlying code that is hosted online, but you will never need to *push/*update what is stored on the tsml GitHub as this will be public)

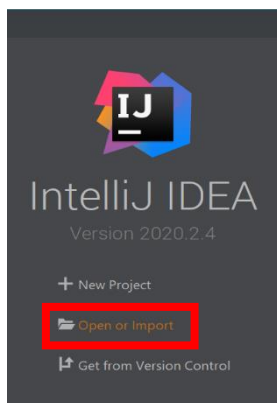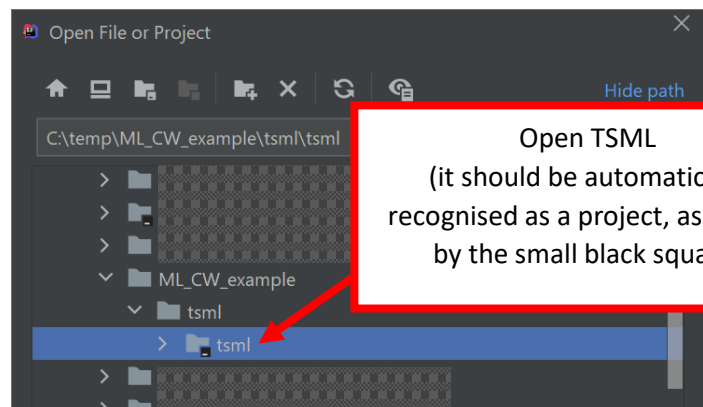**Switch to origin/ml6002b-coursework:**

Before:
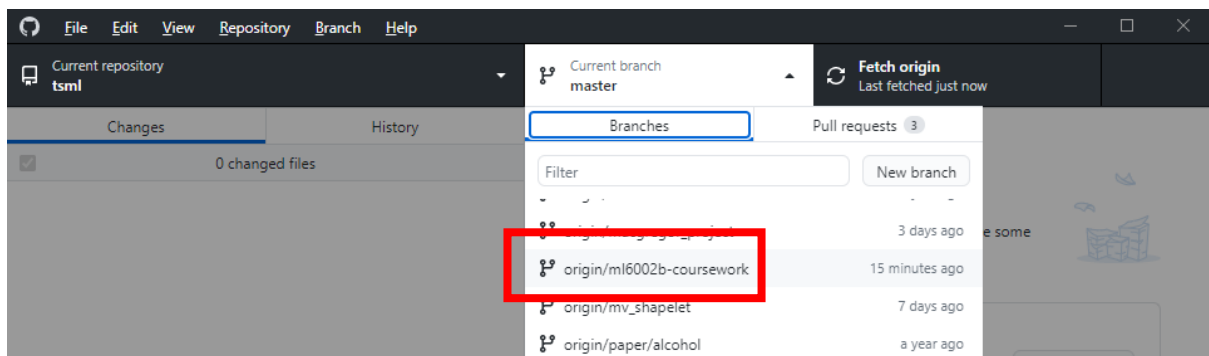
After:
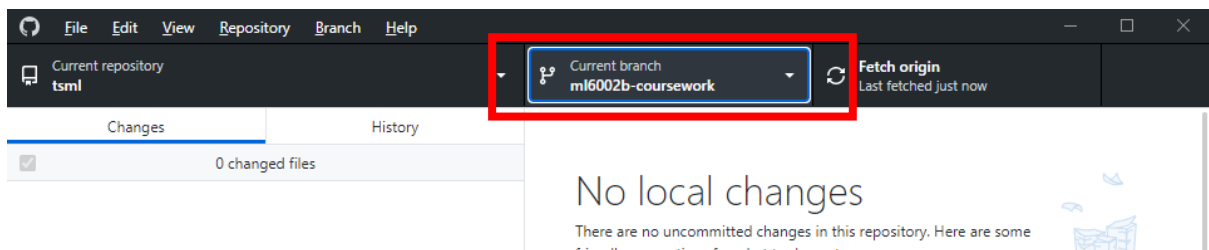
Opening in IntelliJ

The instructions that follow will show you how to get started with the code in *IntelliJ*. You are free to use other IDEs if you prefer but the instructions may be different.
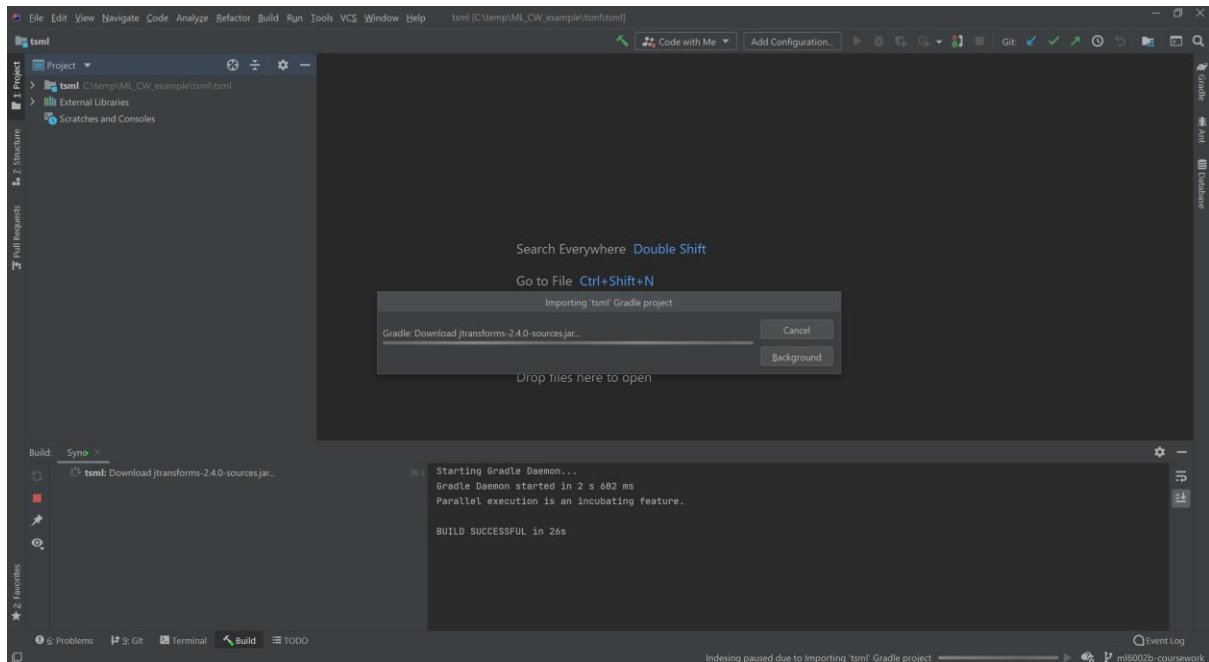

Open a project from the home screen (or File-> Open… if you already have a project open)
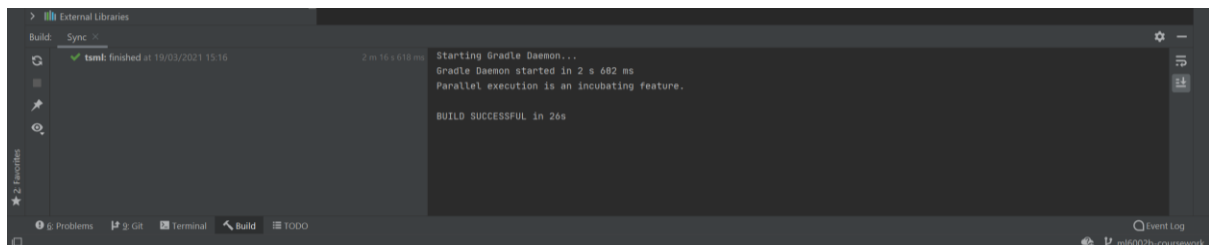
Open TSML
(it should be automatically recognised as a project, as shown by the small black square)

After opening, you may need to wait a minute or two while all of the libraries are downloaded by Gradle – it will look something like this:

Once completed the **Build** prompt will say that the process has finished (it took 2 minutes and 16 seconds for me):


To confirm the installation, it is also a good idea to press the Gradle button and rebuild (this is required *sometimes* but will likely not do anything extra for you)
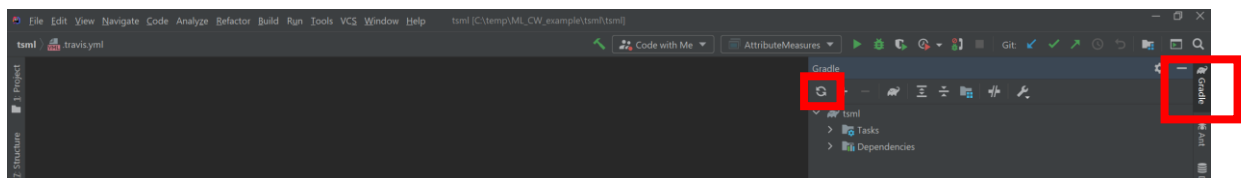
Now you are ready to start coding! (quick example on the final page)

You will find the package that you need for the CW under **src -> main -> java -> ml_6002b_coursework**. To test the installation in the screenshot below I have made an AttributeMeasures class and just included a simple "Hello World!" main method to check the installation.

When you run for the first time the build will be relatively quite slow (approx. 1-2 mins) so do not panic if it takes a while/seems slow. Subsequent runs should be much faster one the initial run has completed as a number of one-off background operations will not be required again.
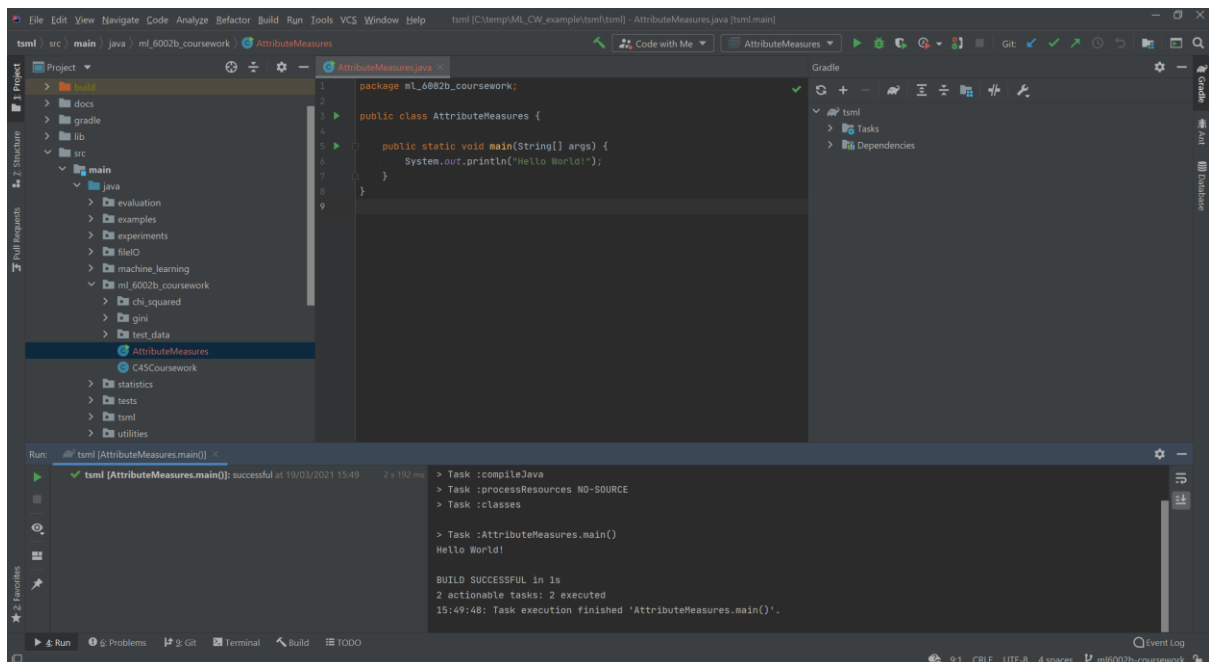
Once you have gotten this far you will now be ready to start implementing the coursework assignment – good luck!