

## Homework #2 (HDFS)

Due: 11:59pm, Friday, February 24, 2023

Points: 100

In this homework, you are asked to emulate the file system structure of HDFS using Firebase and allow the export of its structure in the XML format. The emulated HDFS (or EDFS) should support ls, create, mkdir, rmdir, rm, and export commands as detailed below.

Your task is to write a Python script edfs.py to support the above commands.

Formats:

- `python edfs.py -ls <dir>`

where <dir> is a directory name in EDFS. The command should return names of all files or directories under the given director <dir>. For example, “`edfs.py -ls /user`” will list all files or directories under /user. For another example, “`edfs.py -ls /`” will list all files or directories under the EDFS root directory.

- `python edfs.py -mkdir <path>`

where <path> is a path expression that starts with /, e.g., “`-mkdir /user/dsci551`” will create a directory dsci551 under /user. Note if “/user” does not exist yet, the command will create it too. The script should report if the command was executed successfully or not.

Note there is no need to create a root directory. So the command “`-mkdir /`” should report that / already exists.

- `python edfs.py -rmdir <path>`  
which will remove the directory <path> when the path is not empty, otherwise, it will report an error.
- `python edfs.py -rm <path-to-file>`  
which will remove a file, e.g., /user/dsci551/hello.txt and report errors if the file does not exist or the path points to a directory instead of a file.
- `python edfs.py -create <path-to-file>`  
which will create a file in the specified path. Note the file is empty and your implementation only needs to add the file name to your emulated file system. No need to add data for the file. The command should report an error if the file already exists.
- `python edfs.py -export output.xml`

which will output the EDFS structure in XML format. Note an empty XML element (e.g., `<hello1.txt/>`, represents a file in EDFS. The format of XML is as follows:

```
<root>
  <user>
    <dsci551>
      <hello1.txt/>
      <hello2.txt/>
      <hw1>
        <data1.csv/>
      </hw1>
    </dsci551>
  </user>
</root>
```

**Submissions:** Permitted python libraries are - base64, json, requests, sys, uuid, fileinput.

Please also ensure that your file can be executed as mentioned in the questions. For example:  
`python edfs.py -ls <dir>`

Further, You should assume that the directory names only contain alphanumeric characters and do not contain `"."`. The filenames however, may contain `"."`

Please remember to hard-code your firebase url in your code. Please do not download the entire firebase data to implement commands, instead you should rely on Firebase's functions by sending the appropriate requests.

Submit your edfs.py file via D2L.