MD5 Extractor

The MD5 Extractor is a tool created in Java to extract information from a binary and import that information into a SQL file so it can be used in a database. The files in the database are imported as unknown files, which can then be determined as either good or bad.

Information Collection

The information collected is the MD5 Hash, Firstbytes, File Name & Path, Size, Determination and Date. A folder is automatically created with the name of the MD5 Hash containing text files related to the executable. An SQL file is created which contains the gathered information about the executable which can be uploaded into a database.

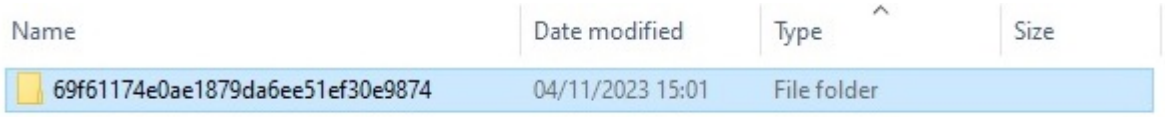
Executing the Program

When the program is run, you are asked for the full path and name of the executable you want to extract information from. In this case it is **test files/test.exe.**

```
Please enter the fullpath of the file
test files/test.exe
Extension = exe
Filename = test
Path = test files
The full path and file name: test files\test.exe
Folder and file name exists: test files\test.exe

The Md5 of the whole file:
69f61174e0ae1879da6ee51ef30e9874
The folder 69f61174e0ae1879da6ee51ef30e9874 has been created:


The firsbytes text file has been created
First 200 bytes of the file one read() at a time
MP 02  00  04  0F ÿ 00  ¸ 00  00  00 @ 1A  00  00  00  00  00  00  00  00
$ 00  00  00  00  00  00  00  00  00  00  00  00  00  00  00  00  00  00

The first200bytes Md5 of the file:
d41d8cd98f00b204e9800998ecf8427e


Total Available Bytes of the whole file: 2126336


Current date is  4/11/2023
Current time is  0:7:49


Text has been outputed to file
```

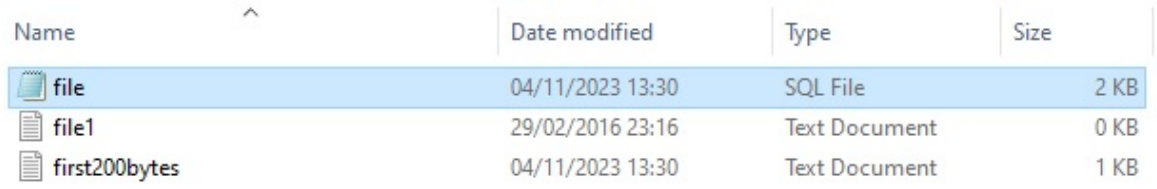
As can be seen from the console output above, information has been extracted from the file and outputted to the screen.

A folder has been created and given the name of the MD5 hash of the file.



Name	Date modified	Type	Size
69f61174e0ae1879da6ee51ef30e9874	04/11/2023 15:01	File folder	

Inside this folder three files have been created.



Name	Date modified	Type	Size
file	04/11/2023 13:30	SQL File	2 KB
file1	29/02/2016 23:16	Text Document	0 KB
first200bytes	04/11/2023 13:30	Text Document	1 KB

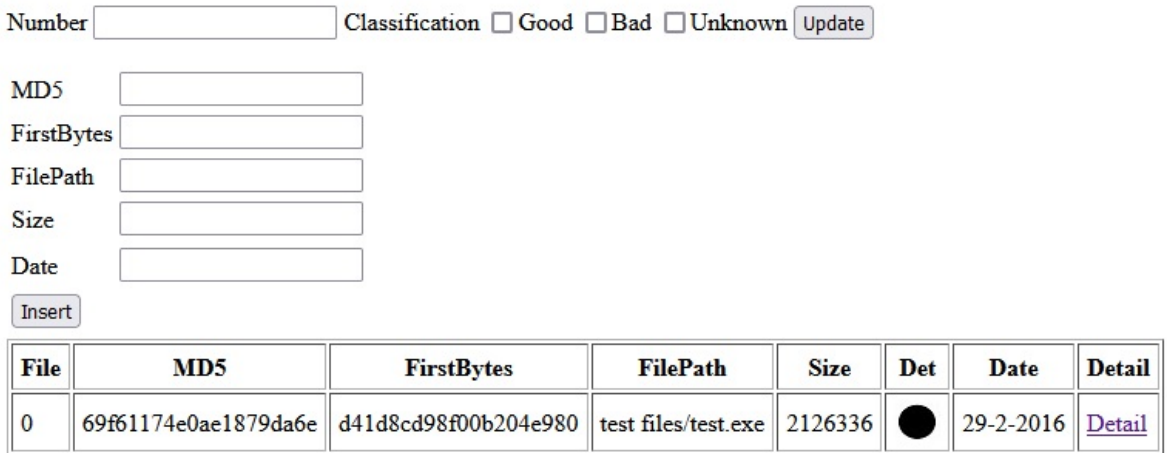
The first file is the SQL file, the second is a file called file1 and the third is a file called first200bytes. These files will be used in my databases or other projects.

The SQL file, when inspected shows the information extracted from the MD5.

```
VALUES(1,'69f61174e0ae1879da6ee51ef30e9874','d41d8cd98f00b204e9800998ecf8427e',' test files/test.exe ',2126336,'unknown.jpeg',
```

The Database

When this SQL File is imported into the database it looks like this.



Number Classification Good Bad Unknown Update

MD5
FirstBytes
FilePath
Size
Date
Insert

File	MD5	FirstBytes	FilePath	Size	Det	Date	Detail
0	69f61174e0ae1879da6e	d41d8cd98f00b204e980	test files/test.exe	2126336	●	29-2-2016	Detail

Here I was able to update the determination value from unknown to bad.