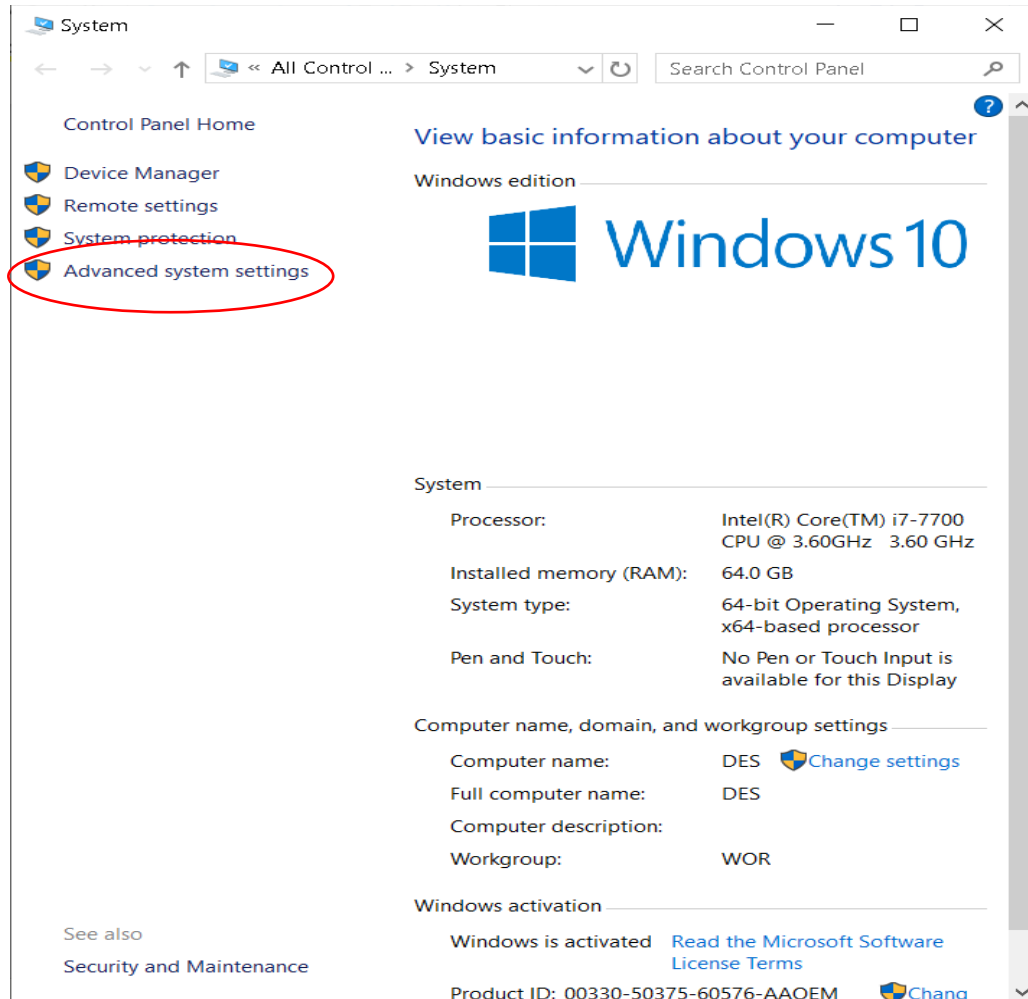
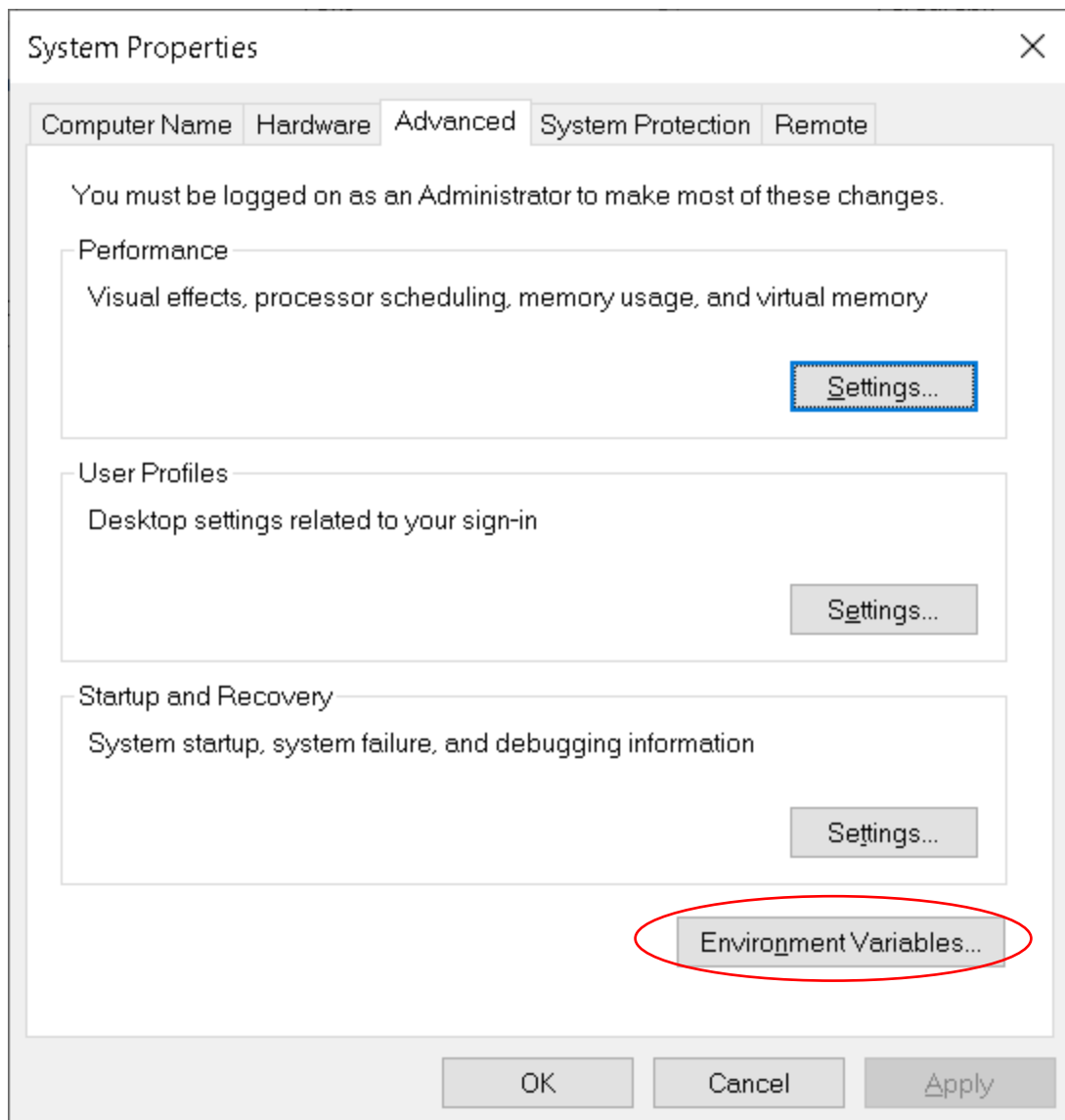


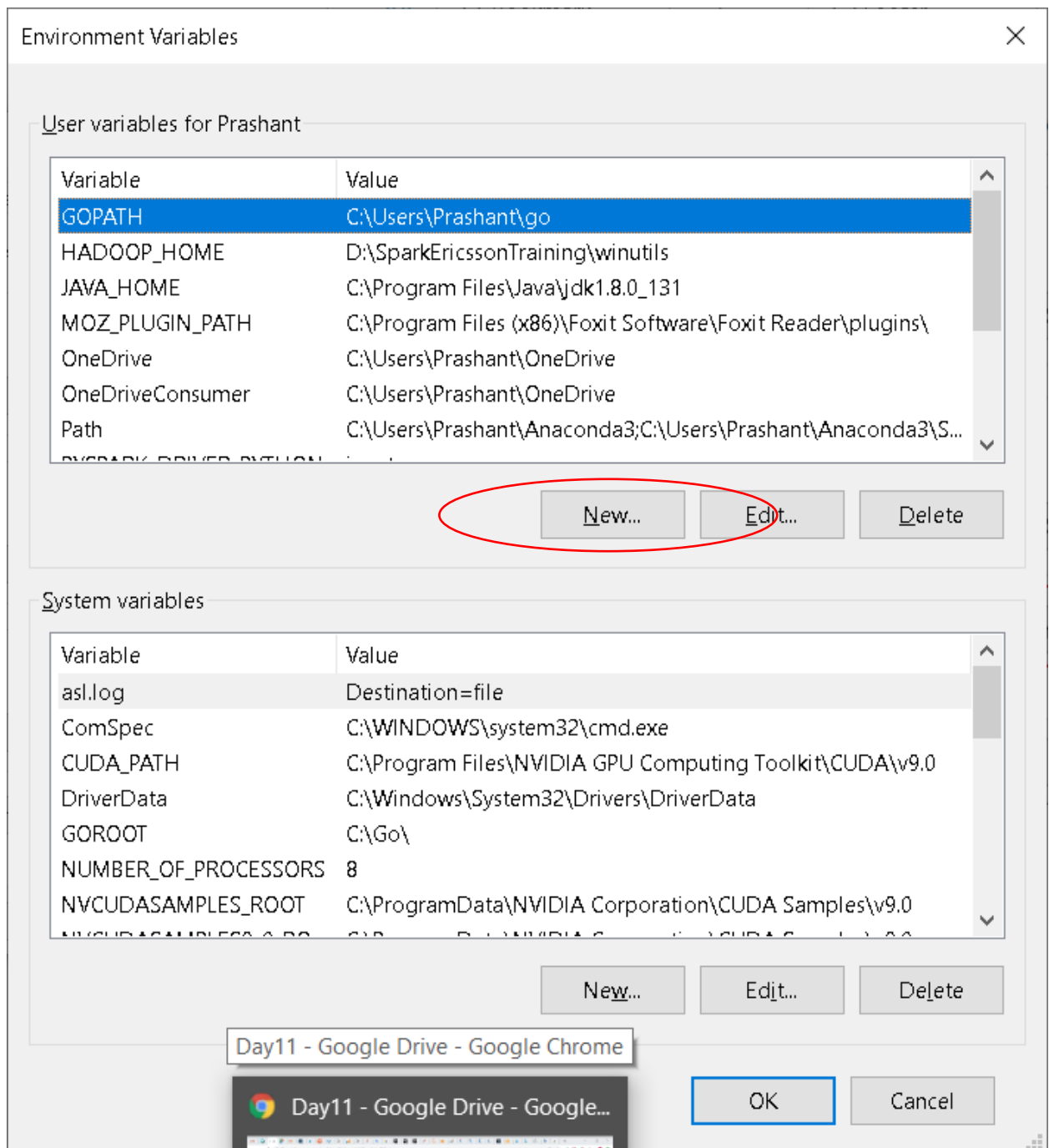
Instructions to enable Apache Spark in Anaconda Installation

1. Download Apache Spark tar file and extract the same
2. Setup the following environment variables

Control Panel > System > Advanced System settings







You will need to create new User Variable as shown

VARIABLE NAME	VARIABLE VALUE
HADOOP_HOME	E:\DataScienceSL\winutils (assuming winutils.exe is in this location. You will find the same in the drive link)
JAVA_HOME	C:\Program Files\Java\jdk1.8.0_131
PYSPARK_DRIVER_PYTHON	jupyter
PYSPARK_DRIVER_PYTHON_OPTS	Notebook

SPARK_HOME	E:\DataScienceSL\MachineLearningSection\Day10\spark-2.4.2-bin-hadoop2.7
------------	---

Note winutils.exe must be stored in directory as shown below

E:\DataScienceSL\winutils\bin\winutils.exe

3. Start Anaconda Prompt. Navigate to the location where you have extracted the Spark folder and type the following command to open Jupyter Notebook enabled with Spark libraries

```
Administrator: Anaconda Prompt

(base) C:\WINDOWS\system32>e:

(base) E:\>cd DataScienceSL

(base) E:\DataScienceSL>cd MachineLearningSection

(base) E:\DataScienceSL\MachineLearningSection>jupyter notebook_
```