**Homework Lab 5.1 Logistic Regression (29/4/2022)**

ລະຫັດນັກສຶກສາ: 205N0083.19

ຊື່ ແລະ ນາມສະກຸນ: ທ ເຊັງວ່າງ ບຼົ່ງໄມ 3cw1

ຈົ່ງຕອບຄຳຖາມຕໍ່ໄປນີ້ໃຫ້ສຳເລັດດ້ວຍການນຳໃຊ້ຄຳສັ່ງຂອງ Python:

1.ຈາກຊຸດຂໍ້ມູນ (Data Set) Social\_Network\_Ads.csv, ຈົ່ງບອກຈຳນວນຖັນ, ແຖວ (shape) ແລະ ເພີ່ມຂໍ້ມູນໃນຕາຕະລາງລຸ່ມນີ້ໃຫ້ສຳເລັດ:

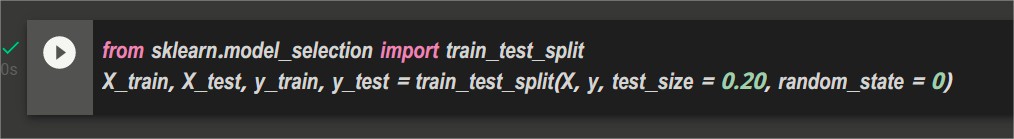
|  |  |
| --- | --- |
| ຖັນ | 5 |
| ແຖວ | 400 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | User ID | Gender | Age | EstimatedSalary | **Purchased** |
|  | 15624510 | Male | 19 | 19000 | 0 |
|  | 15810944 | Male | 35 | 20000 | 0 |
|  | 15668575 | Female | 26 | 43000 | 0 |

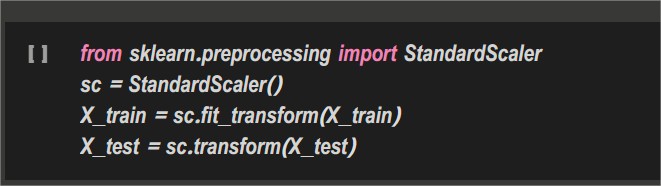
2. ຈົ່ງກຳນົດຕົວປ່ຽນເອກະລາດ (Independent Variables X) ໃຫ້ເປັນ User ID ແລະ EstimatedSalary. ກຳນົດຕົວປ່ຽນຕາມ (Dependent Variables y) ໃຫ້ເປັນ Purchased.



3. ຫຼັງຈາກຕຽມຊຸດຂໍ້ມູນສຳເລັດ, ຈົ່ງຂຽນຄຳສັ່ງເພື່ອແຍກຊຸດຂໍ້ມູນອອກເປັນສອງພາກສ່ວນຄື: ຊຸດຮຽນ 80% ແລະ ຊຸດທົດສອບ 20%?



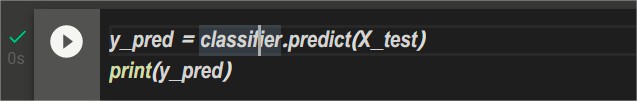
4. ຈົ່ງທຳການຕຽມຊຸດຂໍ້ມູນ (Preprocessing) ດ້ວຍການເຮັດ StandardScaler ຂອງຕົວປ່ຽນອິດສະຫຼະ X\_trainແລະ X\_test.



5. ຈົ່ງສ້າງໂມເດວ Logistic Regression ແລະ ທຳການປະມວນຜົນ (fit) ຊຸດຂໍ້ມູນຮຽນຈາກຂໍ້ 4



6. ຈົ່ງທົດສອບໂມເດວດ້ວຍການpredict(X\_test).

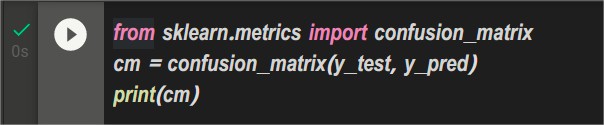


[0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0 0 1 0 0 1 0 1 0 1 0 0 0 0 1 0 0 0 0 0 0

0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0

0 0 0 0 0 1]

7. ຈົ່ງທຳການprocessing ດ້ວຍconfusion\_matrix

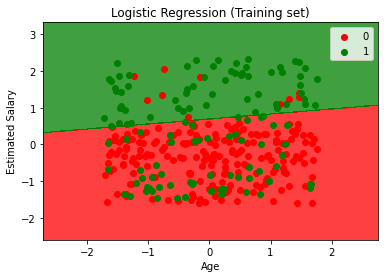


[[55 3]

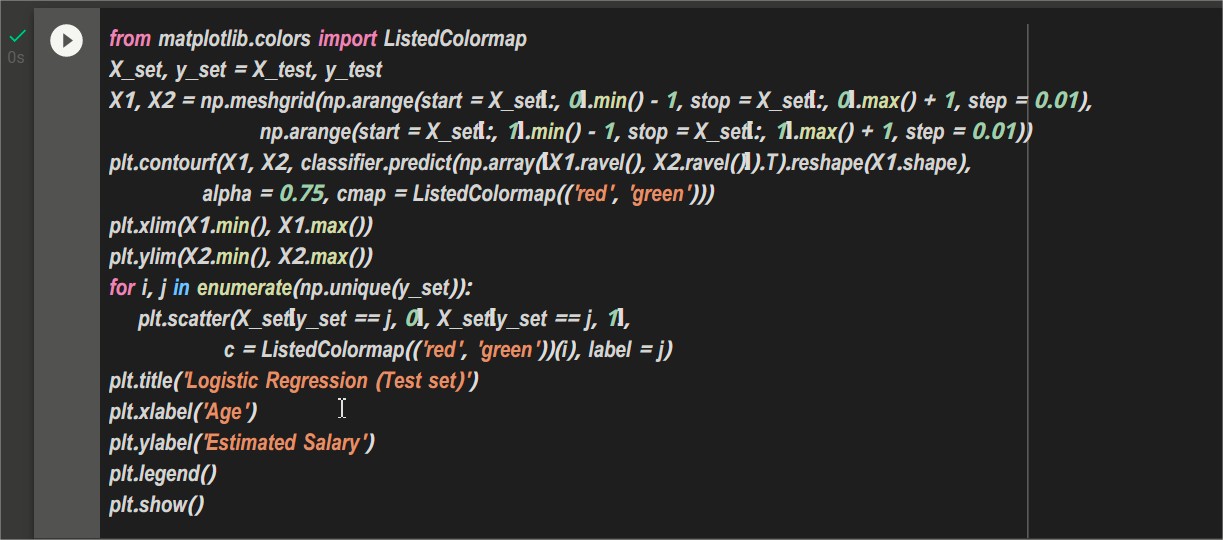
[12 10]]

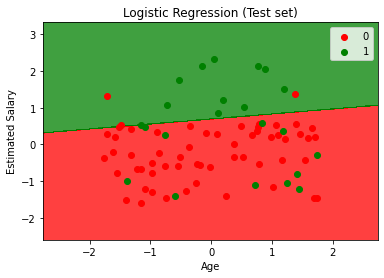
8.ຈົ່ງສະແດງຂໍ້ມູນຊຸດຮຽນ (X\_train, y\_train) ດ້ວຍGraph ບົນພື້ນຖານຊຸດຄຳສັ່ງ matplotlib.





9.ຈົ່ງສະແດງຂໍ້ມູນຊຸດຮຽນ (X\_test, y\_test) ດ້ວຍGraph ບົນພື້ນຖານຊຸດຄຳສັ່ງ matplotlib.

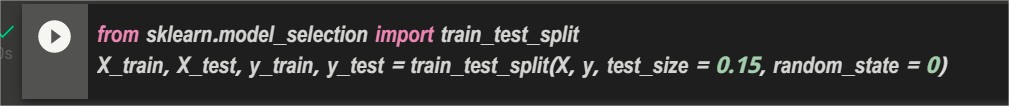




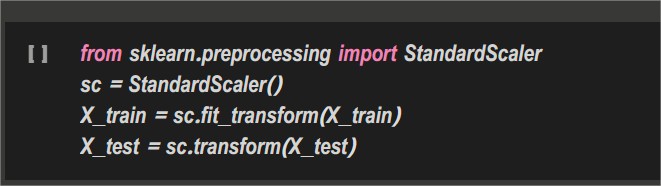
10. ຈົ່ງກຳນົດຕົວປ່ຽນເອກະລາດ (Independent Variables X) ໃຫ້ເປັນ User ID, Ageແລະ EstimatedSalary. ກຳນົດຕົວປ່ຽນຕາມ (Dependent Variables y) ໃຫ້ເປັນ Purchased.



11. ຫຼັງຈາກຕຽມຊຸດຂໍ້ມູນສຳເລັດ, ຈົ່ງຂຽນຄຳສັ່ງເພື່ອແຍກຊຸດຂໍ້ມູນອອກເປັນສອງພາກສ່ວນຄື: ຊຸດຮຽນ 85% ແລະ ຊຸດທົດສອບ 15%?



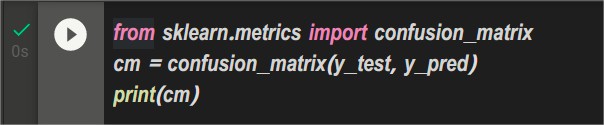
12. ຈົ່ງທຳການຕຽມຊຸດຂໍ້ມູນ (Preprocessing) ດ້ວຍການເຮັດ StandardScaler ຂອງຕົວປ່ຽນອິດສະຫຼະ X\_trainແລະ X\_test.



13. ຈົ່ງສ້າງໂມເດວ Logistic Regression ແລະ ທຳການປະມວນຜົນ (fit) ຊຸດຂໍ້ມູນຮຽນຈາກຂໍ້ 1.12



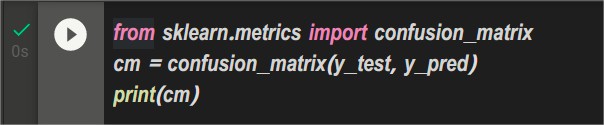
14. ຈົ່ງທົດສອບໂມເດວດ້ວຍການpredict(X\_test).



[0 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 1 0 0 1 0 1 0 1 0 0 0 0 0 0 1 0 0 0 0

0 0 1 0 0 0 0 1 0 0 1 0 1 1 0 0 0 1 0 0 0 0 0]

15 ຈົ່ງທຳການprocessing ດ້ວຍconfusion\_matrix



[[44 1]

[ 3 12]]