First AI predictor

April 4, 2024

[1]:
```
!pip install "tqdm == 4.60.0"
!pip install "seaborn == 0.9.0"
!pip install "skillsnetwork == 0.20.6"
!pip install "scikit-plot == 0.3.7"
!pip install "numpy == 1.19.5"
```

```
Requirement already satisfied: tqdm==4.60.0 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (4.60.0)
Requirement already satisfied: seaborn==0.9.0 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (0.9.0)
Requirement already satisfied: matplotlib>=1.4.3 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
seaborn==0.9.0) (3.5.3)
Requirement already satisfied: numpy>=1.9.3 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
seaborn==0.9.0) (1.21.6)
Requirement already satisfied: pandas>=0.15.2 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
seaborn==0.9.0) (1.3.5)
Requirement already satisfied: scipy>=0.14.0 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
seaborn==0.9.0) (1.7.3)
Requirement already satisfied: cycler>=0.10 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
matplotlib>=1.4.3->seaborn==0.9.0) (0.11.0)
Requirement already satisfied: fonttools>=4.22.0 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
matplotlib>=1.4.3->seaborn==0.9.0) (4.38.0)
Requirement already satisfied: kiwisolver>=1.0.1 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
matplotlib>=1.4.3->seaborn==0.9.0) (1.4.4)
Requirement already satisfied: packaging>=20.0 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
matplotlib>=1.4.3->seaborn==0.9.0) (23.1)
Requirement already satisfied: pillow>=6.2.0 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
matplotlib>=1.4.3->seaborn==0.9.0) (8.1.0)
Requirement already satisfied: pyparsing>=2.2.1 in
```

```
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
matplotlib>=1.4.3->seaborn==0.9.0) (3.0.9)
Requirement already satisfied: python-dateutil>=2.7 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
matplotlib>=1.4.3->seaborn==0.9.0) (2.8.2)
Requirement already satisfied: pytz>=2017.3 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
pandas>=0.15.2->seaborn==0.9.0) (2023.3)
Requirement already satisfied: typing-extensions in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
kiwisolver>=1.0.1->matplotlib>=1.4.3->seaborn==0.9.0) (4.5.0)
Requirement already satisfied: six>=1.5 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from python-
dateutil>=2.7->matplotlib>=1.4.3->seaborn==0.9.0) (1.16.0)
Requirement already satisfied: skillsnetwork==0.20.6 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (0.20.6)
Requirement already satisfied: ipython in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
skillsnetwork==0.20.6) (7.33.0)
Requirement already satisfied: ipywidgets<8,>=7 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
skillsnetwork==0.20.6) (7.4.2)
Requirement already satisfied: requests<3,>=2 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
skillsnetwork==0.20.6) (2.29.0)
Requirement already satisfied: tqdm<5,>=4 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
skillsnetwork==0.20.6) (4.60.0)
Requirement already satisfied: ipykernel>=4.5.1 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
ipywidgets<8,>=7->skillsnetwork==0.20.6) (6.16.2)
Requirement already satisfied: traitlets>=4.3.1 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
ipywidgets<8,>=7->skillsnetwork==0.20.6) (5.9.0)
Requirement already satisfied: nbformat>=4.2.0 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
ipywidgets<8,>=7->skillsnetwork==0.20.6) (5.8.0)
Requirement already satisfied: widgetsnbextension~=3.4.0 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
ipywidgets<8,>=7->skillsnetwork==0.20.6) (3.4.2)
Requirement already satisfied: setuptools>=18.5 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
ipython->skillsnetwork==0.20.6) (67.7.2)
Requirement already satisfied: jedi>=0.16 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
ipython->skillsnetwork==0.20.6) (0.18.2)
Requirement already satisfied: decorator in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
```

```
ipython->skillsnetwork==0.20.6) (5.1.1)
Requirement already satisfied: pickleshare in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
ipython->skillsnetwork==0.20.6) (0.7.5)
Requirement already satisfied: prompt-toolkit!=3.0.0,!=3.0.1,<3.1.0,>=2.0.0 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
ipython->skillsnetwork==0.20.6) (3.0.38)
Requirement already satisfied: pygments in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
ipython->skillsnetwork==0.20.6) (2.15.1)
Requirement already satisfied: backcall in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
ipython->skillsnetwork==0.20.6) (0.2.0)
Requirement already satisfied: matplotlib-inline in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
ipython->skillsnetwork==0.20.6) (0.1.6)
Requirement already satisfied: pexpect>4.3 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
ipython->skillsnetwork==0.20.6) (4.8.0)
Requirement already satisfied: charset-normalizer<4,>=2 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
requests<3,>=2->skillsnetwork==0.20.6) (3.1.0)
Requirement already satisfied: idna<4,>=2.5 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
requests<3,>=2->skillsnetwork==0.20.6) (3.4)
Requirement already satisfied: urllib3<1.27,>=1.21.1 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
requests<3,>=2->skillsnetwork==0.20.6) (1.26.15)
Requirement already satisfied: certifi>=2017.4.17 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
requests<3,>=2->skillsnetwork==0.20.6) (2023.5.7)
Requirement already satisfied: debugpy>=1.0 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
ipykernel>=4.5.1->ipywidgets<8,>=7->skillsnetwork==0.20.6) (1.6.3)
Requirement already satisfied: jupyter-client>=6.1.12 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
ipykernel>=4.5.1->ipywidgets<8,>=7->skillsnetwork==0.20.6) (7.4.9)
Requirement already satisfied: nest-asyncio in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
ipykernel>=4.5.1->ipywidgets<8,>=7->skillsnetwork==0.20.6) (1.5.6)
Requirement already satisfied: packaging in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
ipykernel>=4.5.1->ipywidgets<8,>=7->skillsnetwork==0.20.6) (23.1)
Requirement already satisfied: psutil in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
ipykernel>=4.5.1->ipywidgets<8,>=7->skillsnetwork==0.20.6) (5.9.3)
Requirement already satisfied: pyzmq>=17 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
```

```
ipykernel>=4.5.1->ipywidgets<8,>=7->skillsnetwork==0.20.6) (24.0.1)
Requirement already satisfied: tornado>=6.1 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
ipykernel>=4.5.1->ipywidgets<8,>=7->skillsnetwork==0.20.6) (6.2)
Requirement already satisfied: parso<0.9.0,>=0.8.0 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
jedi>=0.16->ipython->skillsnetwork==0.20.6) (0.8.3)
Requirement already satisfied: fastjsonschema in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
nbformat>=4.2.0->ipywidgets<8,>=7->skillsnetwork==0.20.6) (2.16.3)
Requirement already satisfied: importlib-metadata>=3.6 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
nbformat>=4.2.0->ipywidgets<8,>=7->skillsnetwork==0.20.6) (4.11.4)
Requirement already satisfied: jsonschema>=2.6 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
nbformat>=4.2.0->ipywidgets<8,>=7->skillsnetwork==0.20.6) (4.17.3)
Requirement already satisfied: jupyter-core in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
nbformat>=4.2.0->ipywidgets<8,>=7->skillsnetwork==0.20.6) (4.12.0)
Requirement already satisfied: ptyprocess>=0.5 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
pexpect>4.3->ipython->skillsnetwork==0.20.6) (0.7.0)
Requirement already satisfied: wcwidth in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from prompt-
toolkit!=3.0.0,!=3.0.1,<3.1.0,>=2.0.0->ipython->skillsnetwork==0.20.6) (0.2.6)
Requirement already satisfied: notebook>=4.4.1 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
widgetsnbextension~=3.4.0->ipywidgets<8,>=7->skillsnetwork==0.20.6) (6.5.4)
Requirement already satisfied: zipp>=0.5 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from importlib-
metadata>=3.6->nbformat>=4.2.0->ipywidgets<8,>=7->skillsnetwork==0.20.6)
(3.15.0)
Requirement already satisfied: typing-extensions>=3.6.4 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from importlib-
metadata>=3.6->nbformat>=4.2.0->ipywidgets<8,>=7->skillsnetwork==0.20.6) (4.5.0)
Requirement already satisfied: attrs>=17.4.0 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
jsonschema>=2.6->nbformat>=4.2.0->ipywidgets<8,>=7->skillsnetwork==0.20.6)
(23.1.0)
Requirement already satisfied: importlib-resources>=1.4.0 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
jsonschema>=2.6->nbformat>=4.2.0->ipywidgets<8,>=7->skillsnetwork==0.20.6)
(5.12.0)
Requirement already satisfied: pkgutil-resolve-name>=1.3.10 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
jsonschema>=2.6->nbformat>=4.2.0->ipywidgets<8,>=7->skillsnetwork==0.20.6)
(1.3.10)
Requirement already satisfied: pyrsistent!=0.17.0,!=0.17.1,!=0.17.2,>=0.14.0 in
```

```
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
jsonschema>=2.6->nbformat>=4.2.0->ipywidgets<8,>=7->skillsnetwork==0.20.6)
(0.19.3)
Requirement already satisfied: entrypoints in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from jupyter-
client>=6.1.12->ipykernel>=4.5.1->ipywidgets<8,>=7->skillsnetwork==0.20.6) (0.4)
Requirement already satisfied: python-dateutil>=2.8.2 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from jupyter-
client>=6.1.12->ipykernel>=4.5.1->ipywidgets<8,>=7->skillsnetwork==0.20.6)
(2.8.2)
Requirement already satisfied: jinja2 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from notebook>=4
.4.1->widgetsnbextension~=3.4.0->ipywidgets<8,>=7->skillsnetwork==0.20.6)
(3.1.2)
Requirement already satisfied: argon2-cffi in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from notebook>=4
.4.1->widgetsnbextension~=3.4.0->ipywidgets<8,>=7->skillsnetwork==0.20.6)
(21.3.0)
Requirement already satisfied: ipython-genutils in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from notebook>=4
.4.1->widgetsnbextension~=3.4.0->ipywidgets<8,>=7->skillsnetwork==0.20.6)
(0.2.0)
Requirement already satisfied: nbconvert>=5 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from notebook>=4
.4.1->widgetsnbextension~=3.4.0->ipywidgets<8,>=7->skillsnetwork==0.20.6)
(7.4.0)
Requirement already satisfied: Send2Trash>=1.8.0 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from notebook>=4
.4.1->widgetsnbextension~=3.4.0->ipywidgets<8,>=7->skillsnetwork==0.20.6)
(1.8.2)
Requirement already satisfied: terminado>=0.8.3 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from notebook>=4
.4.1->widgetsnbextension~=3.4.0->ipywidgets<8,>=7->skillsnetwork==0.20.6)
(0.17.1)
Requirement already satisfied: prometheus-client in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from notebook>=4
.4.1->widgetsnbextension~=3.4.0->ipywidgets<8,>=7->skillsnetwork==0.20.6)
(0.16.0)
Requirement already satisfied: nbclassic>=0.4.7 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from notebook>=4
.4.1->widgetsnbextension~=3.4.0->ipywidgets<8,>=7->skillsnetwork==0.20.6)
(1.0.0)
Requirement already satisfied: jupyter-server>=1.8 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from nbclassic>=
0.4.7->notebook>=4.4.1->widgetsnbextension~=3.4.0->ipywidgets<8,>=7->skillsnetwo
rk==0.20.6) (1.24.0)
Requirement already satisfied: notebook-shim>=0.2.3 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from nbclassic>=
```

```
0.4.7->notebook>=4.4.1->widgetsnbextension~=3.4.0->ipywidgets<8,>=7->skillsnetwo
rk==0.20.6) (0.2.3)
Requirement already satisfied: beautifulsoup4 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from nbconvert>=
5->notebook>=4.4.1->widgetsnbextension~=3.4.0->ipywidgets<8,>=7->skillsnetwork==
0.20.6) (4.11.1)
Requirement already satisfied: bleach in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from nbconvert>=
5->notebook>=4.4.1->widgetsnbextension~=3.4.0->ipywidgets<8,>=7->skillsnetwork==
0.20.6) (6.0.0)
Requirement already satisfied: defusedxml in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from nbconvert>=
5->notebook>=4.4.1->widgetsnbextension~=3.4.0->ipywidgets<8,>=7->skillsnetwork==
0.20.6) (0.7.1)
Requirement already satisfied: jupyterlab-pygments in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from nbconvert>=
5->notebook>=4.4.1->widgetsnbextension~=3.4.0->ipywidgets<8,>=7->skillsnetwork==
0.20.6) (0.2.2)
Requirement already satisfied: markupsafe>=2.0 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from nbconvert>=
5->notebook>=4.4.1->widgetsnbextension~=3.4.0->ipywidgets<8,>=7->skillsnetwork==
0.20.6) (2.1.1)
Requirement already satisfied: mistune<3,>=2.0.3 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from nbconvert>=
5->notebook>=4.4.1->widgetsnbextension~=3.4.0->ipywidgets<8,>=7->skillsnetwork==
0.20.6) (2.0.5)
Requirement already satisfied: nbclient>=0.5.0 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from nbconvert>=
5->notebook>=4.4.1->widgetsnbextension~=3.4.0->ipywidgets<8,>=7->skillsnetwork==
0.20.6) (0.7.4)
Requirement already satisfied: pandocfilters>=1.4.1 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from nbconvert>=
5->notebook>=4.4.1->widgetsnbextension~=3.4.0->ipywidgets<8,>=7->skillsnetwork==
0.20.6) (1.5.0)
Requirement already satisfied: tinycss2 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from nbconvert>=
5->notebook>=4.4.1->widgetsnbextension~=3.4.0->ipywidgets<8,>=7->skillsnetwork==
0.20.6) (1.2.1)
Requirement already satisfied: six>=1.5 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from python-
dateutil>=2.8.2->jupyter-
client>=6.1.12->ipykernel>=4.5.1->ipywidgets<8,>=7->skillsnetwork==0.20.6)
(1.16.0)
Requirement already satisfied: argon2-cffi-bindings in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from argon2-cffi
->notebook>=4.4.1->widgetsnbextension~=3.4.0->ipywidgets<8,>=7->skillsnetwork==0
.20.6) (21.2.0)
Requirement already satisfied: anyio<4,>=3.1.0 in
```

```
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from jupyter-ser
ver>=1.8->nbclassic>=0.4.7->notebook>=4.4.1->widgetsnbextension~=3.4.0->ipywidge
ts<8,>=7->skillsnetwork==0.20.6) (3.6.2)
Requirement already satisfied: websocket-client in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from jupyter-ser
ver>=1.8->nbclassic>=0.4.7->notebook>=4.4.1->widgetsnbextension~=3.4.0->ipywidge
ts<8,>=7->skillsnetwork==0.20.6) (1.5.1)
Requirement already satisfied: cffi>=1.0.1 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
argon2-cffi-bindings->argon2-cffi->notebook>=4.4.1->widgetsnbextension~=3.4.0->i
pywidgets<8,>=7->skillsnetwork==0.20.6) (1.15.1)
Requirement already satisfied: soupsieve>1.2 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from beautifulso
up4->nbconvert>=5->notebook>=4.4.1->widgetsnbextension~=3.4.0->ipywidgets<8,>=7-
>skillsnetwork==0.20.6) (2.3.2.post1)
Requirement already satisfied: webencodings in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from bleach->nbc
onvert>=5->notebook>=4.4.1->widgetsnbextension~=3.4.0->ipywidgets<8,>=7->skillsn
etwork==0.20.6) (0.5.1)
Requirement already satisfied: sniffio>=1.1 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
anyio<4,>=3.1.0->jupyter-server>=1.8->nbclassic>=0.4.7->notebook>=4.4.1->widgets
nbextension~=3.4.0->ipywidgets<8,>=7->skillsnetwork==0.20.6) (1.3.0)
Requirement already satisfied: pycparser in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
cffi>=1.0.1->argon2-cffi-bindings->argon2-cffi->notebook>=4.4.1->widgetsnbextens
ion~=3.4.0->ipywidgets<8,>=7->skillsnetwork==0.20.6) (2.21)
Collecting scikit-plot==0.3.7
  Downloading scikit_plot-0.3.7-py3-none-any.whl (33 kB)
Requirement already satisfied: matplotlib>=1.4.0 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from scikit-
plot==0.3.7) (3.5.3)
Requirement already satisfied: scikit-learn>=0.18 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from scikit-
plot==0.3.7) (0.24.0)
Requirement already satisfied: scipy>=0.9 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from scikit-
plot==0.3.7) (1.7.3)
Requirement already satisfied: joblib>=0.10 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from scikit-
plot==0.3.7) (1.3.2)
Requirement already satisfied: cycler>=0.10 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
matplotlib>=1.4.0->scikit-plot==0.3.7) (0.11.0)
Requirement already satisfied: fonttools>=4.22.0 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
matplotlib>=1.4.0->scikit-plot==0.3.7) (4.38.0)
Requirement already satisfied: kiwisolver>=1.0.1 in
```

```
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
matplotlib>=1.4.0->scikit-plot==0.3.7) (1.4.4)
Requirement already satisfied: numpy>=1.17 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
matplotlib>=1.4.0->scikit-plot==0.3.7) (1.21.6)
Requirement already satisfied: packaging>=20.0 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
matplotlib>=1.4.0->scikit-plot==0.3.7) (23.1)
Requirement already satisfied: pillow>=6.2.0 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
matplotlib>=1.4.0->scikit-plot==0.3.7) (8.1.0)
Requirement already satisfied: pyparsing>=2.2.1 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
matplotlib>=1.4.0->scikit-plot==0.3.7) (3.0.9)
Requirement already satisfied: python-dateutil>=2.7 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
matplotlib>=1.4.0->scikit-plot==0.3.7) (2.8.2)
Requirement already satisfied: threadpoolctl>=2.0.0 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from scikit-
learn>=0.18->scikit-plot==0.3.7) (3.1.0)
Requirement already satisfied: typing-extensions in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from
kiwisolver>=1.0.1->matplotlib>=1.4.0->scikit-plot==0.3.7) (4.5.0)
Requirement already satisfied: six>=1.5 in
/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from python-
dateutil>=2.7->matplotlib>=1.4.0->scikit-plot==0.3.7) (1.16.0)
Installing collected packages: scikit-plot
Successfully installed scikit-plot-0.3.7
Collecting numpy==1.19.5
  Downloading numpy-1.19.5-cp37-cp37m-manylinux2010_x86_64.whl (14.8 MB)
                             14.8/14.8 MB
58.0 MB/s eta 0:00:0000:0100:01
Installing collected packages: numpy
  Attempting uninstall: numpy
    Found existing installation: numpy 1.21.6
    Uninstalling numpy-1.21.6:
      Successfully uninstalled numpy-1.21.6
Successfully installed numpy-1.19.5
```

[2]:
```
# You can use this section to suppress warnings generated by your code:
def warn(*args, **kwargs):
    pass
import warnings
warnings.warn = warn
warnings.filterwarnings('ignore', category=DeprecationWarning)
warn()
```

```
from tqdm import tqdm
import skillsnetwork
import numpy as np
import pandas as pd
from itertools import accumulate
import matplotlib.pyplot as plt
import matplotlib.patches as mpatches
import seaborn as sns
import scikitplot as skplt
%matplotlib inline
from sklearn.preprocessing import StandardScaler, OneHotEncoder, LabelEncoder
from sklearn.impute import SimpleImputer
from sklearn.pipeline import Pipeline
from sklearn.compose import ColumnTransformer
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import accuracy_score, classification_report

sns.set_context('notebook')
sns.set_style('white')
```

[3]:
```
await skillsnetwork.download_dataset('https://cf-courses-data.s3.us.
  ↪cloud-object-storage.appdomain.cloud/IBMSkillsNetwork-GPXX056YEN/churn.csv')
df_churn_pd = pd.read_csv('churn.csv')

df_churn_pd.head()
```

```
Downloading churn.csv:   0%|          | 0/164241 [00:00<?, ?it/s]

Saved as 'churn.csv'
```

[3]:

	ID	CHURNRISK	GENDER	STATUS	CHILDREN	ESTINCOME	HOMEOWNER	AGE	\
0	1703	Medium	NaN	M	2	28766.9	N	47	
1	1704	Low	NaN	M	2	91272.2	Y	25	
2	1705	Low	NaN	S	0	73228.3	N	42	
3	1706	High	M	M	1	64792.3	N	52	
4	1707	High	F	S	0	93322.1	Y	40	

	TOTALDOLLARVALUETRADED	TOTALUNITSTRADED	LARGESTSINGLETRANSACTION	\
0	6110.61	58	1527.6525	
1	26992.70	13	13496.3500	
2	22472.25	28	11236.1250	
3	13051.31	36	6525.6550	
4	29922.99	8	14961.4950	

	SMALLESTSINGLETRANSACTION	PERCENTCHANGECALCULATION	DAYSSINCELASTLOGIN	\
0	152.76525	8.70	2	

```
1                1349.63500                      3.25                   4
2                1123.61250                      7.00                   4
3                 652.56550                      9.00                   3
4                1496.14950                      2.00                   4

   DAYSSINCELASTTRADE  NETREALIZEDGAINS_YTD  NETREALIZEDLOSSES_YTD
0                  13                0.0000              152.76525
1                  10             1349.6350                0.00000
2                   5             1123.6125                0.00000
3                   6                0.0000              652.56550
4                   9                0.0000             1496.14950
```

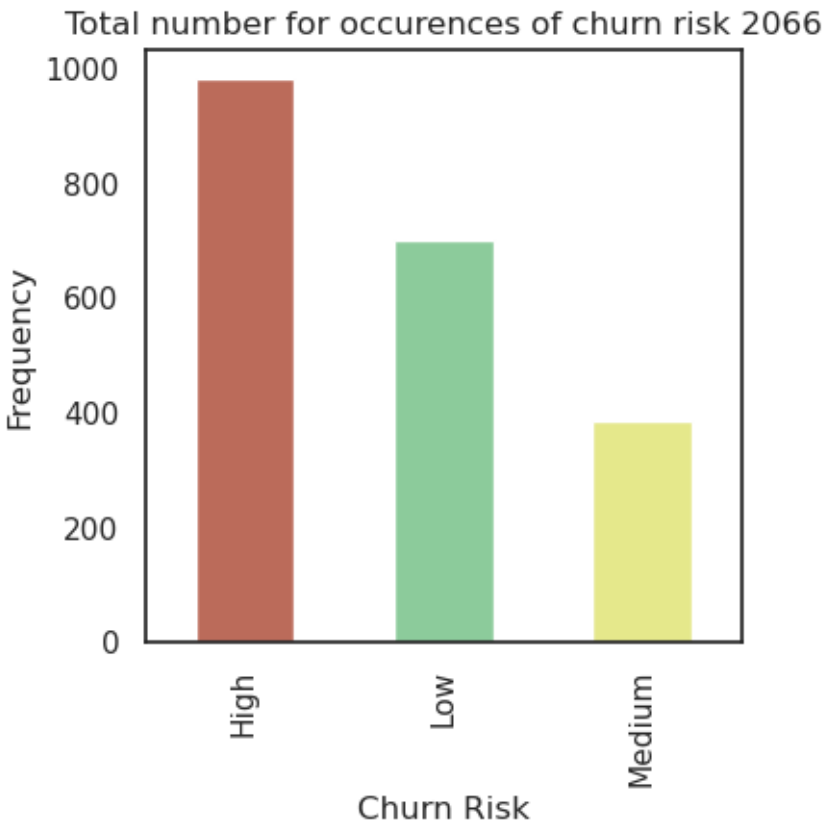
[4]:
```
print("The dataset contains columns of the following data types : \n" +␣
 ↪str(df_churn_pd.dtypes))
```

```
The dataset contains columns of the following data types :
ID                             int64
CHURNRISK                     object
GENDER                        object
STATUS                        object
CHILDREN                       int64
ESTINCOME                    float64
HOMEOWNER                     object
AGE                            int64
TOTALDOLLARVALUETRADED       float64
TOTALUNITSTRADED               int64
LARGESTSINGLETRANSACTION     float64
SMALLESTSINGLETRANSACTION    float64
PERCENTCHANGECALCULATION     float64
DAYSSINCELASTLOGIN             int64
DAYSSINCELASTTRADE             int64
NETREALIZEDGAINS_YTD         float64
NETREALIZEDLOSSES_YTD        float64
dtype: object
```

[5]:
```
print("Each category within the churnrisk column has the following count : ")
print(df_churn_pd.groupby(['CHURNRISK']).size())
#bar chart to show split of data
index = ['High', 'Medium', 'Low']
churn_plot = df_churn_pd['CHURNRISK'].value_counts(sort=True, ascending=False).
 ↪plot(kind='bar',
           figsize=(4,4), title="Total number for occurences of churn risk "
           + str(df_churn_pd['CHURNRISK'].count()), color=['#BB6B5A',␣
 ↪'#8CCB9B', '#E5E88B'])
churn_plot.set_xlabel("Churn Risk")
churn_plot.set_ylabel("Frequency")
```

```
Each category within the churnrisk column has the following count :
CHURNRISK
High      983
Low       699
Medium    384
dtype: int64
```

[5]: `Text(0, 0.5, 'Frequency')`



[6]:
```
# remove columns that are not required
df_churn_pd = df_churn_pd.drop(['ID'], axis=1)
df_churn_pd.head()
```

[6]:

	CHURNRISK	GENDER	STATUS	CHILDREN	ESTINCOME	HOMEOWNER	AGE
0	Medium	NaN	M	2	28766.9	N	47
1	Low	NaN	M	2	91272.2	Y	25
2	Low	NaN	S	0	73228.3	N	42
3	High	M	M	1	64792.3	N	52
4	High	F	S	0	93322.1	Y	40

	TOTALDOLLARVALUETRADED	TOTALUNITSTRADED	LARGESTSINGLETRANSACTION

```
0                6110.61                58                1527.6525
1               26992.70                13               13496.3500
2               22472.25                28               11236.1250
3               13051.31                36                6525.6550
4               29922.99                 8               14961.4950

   SMALLESTSINGLETRANSACTION  PERCENTCHANGECALCULATION  DAYSSINCELASTLOGIN  \
0                  152.76525                      8.70                   2
1                 1349.63500                      3.25                   4
2                 1123.61250                      7.00                   4
3                  652.56550                      9.00                   3
4                 1496.14950                      2.00                   4

   DAYSSINCELASTTRADE  NETREALIZEDGAINS_YTD  NETREALIZEDLOSSES_YTD
0                  13                0.0000              152.76525
1                  10             1349.6350                0.00000
2                   5             1123.6125                0.00000
3                   6                0.0000              652.56550
4                   9                0.0000             1496.14950
```

[7]:
```
# Defining the categorical columns
categoricalColumns = ['GENDER', 'STATUS', 'HOMEOWNER']

print("Categorical columns : ")
print(categoricalColumns)

impute_categorical = SimpleImputer(strategy="most_frequent")

onehot_categorical = OneHotEncoder(handle_unknown='ignore')

categorical_transformer = Pipeline(steps=[('impute', impute_categorical),␣
 ↪('onehot', onehot_categorical)])
```

```
Categorical columns :
['GENDER', 'STATUS', 'HOMEOWNER']
```

[8]:
```
# Defining the numerical columns
numericalColumns = df_churn_pd.select_dtypes(include=[np.float,np.int]).columns

print("Numerical columns : " )
print(numericalColumns)

scaler_numerical = StandardScaler()

numerical_transformer = Pipeline(steps=[('scale', scaler_numerical)])
```

```
Numerical columns :
Index(['CHILDREN', 'ESTINCOME', 'AGE', 'TOTALDOLLARVALUETRADED',
```

```
       'TOTALUNITSTRADED', 'LARGESTSINGLETRANSACTION',
       'SMALLESTSINGLETRANSACTION', 'PERCENTCHANGECALCULATION',
       'DAYSSINCELASTLOGIN', 'DAYSSINCELASTTRADE', 'NETREALIZEDGAINS_YTD',
       'NETREALIZEDLOSSES_YTD'],
      dtype='object')
```

[9]:
```
preprocessorForCategoricalColumns = ColumnTransformer(transformers=[('cat',␣
 ↪categorical_transformer,
                                                                     ␣
 ↪categoricalColumns) ],
                                                remainder="passthrough" )
preprocessorForAllColumns = ColumnTransformer(transformers=[('cat',␣
 ↪categorical_transformer, categoricalColumns),
                                                                ('num',␣
 ↪numerical_transformer, numericalColumns) ],
                                                    remainder="passthrough" )

#. The transformation happens in the pipeline. Temporarily done here to show␣
 ↪what intermediate value looks like
df_churn_pd_temp = preprocessorForCategoricalColumns.fit_transform(df_churn_pd)

print("Data after transforming :")

print(df_churn_pd_temp)

df_churn_pd_temp_2 = preprocessorForAllColumns.fit_transform(df_churn_pd)
print("Data after transforming :")
print(df_churn_pd_temp_2)
```

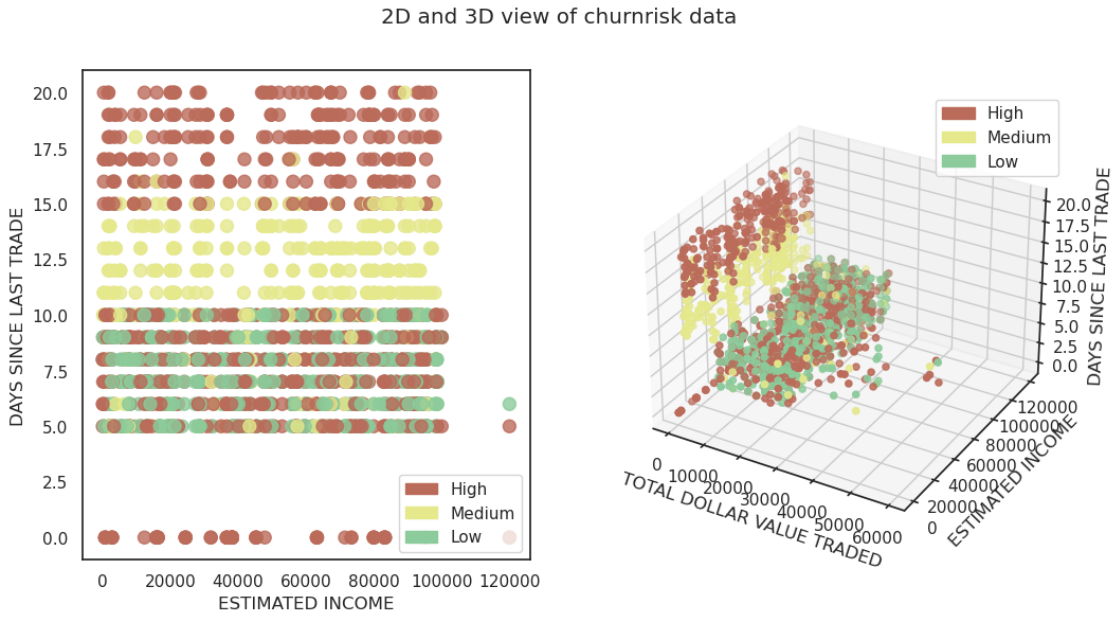
```
Data after transforming :
[[1.0 0.0 0.0 … 13 0.0 152.76525]
 [1.0 0.0 0.0 … 10 1349.635 0.0]
 [1.0 0.0 0.0 … 5 1123.6125 0.0]
 …
 [1.0 0.0 0.0 … 11 0.0 82.50825]
 [1.0 0.0 0.0 … 5 930.093 0.0]
 [0.0 1.0 0.0 … 8 1041.604 0.0]]
Data after transforming :
[[1.0 0.0 0.0 … -0.6792731303453047 -0.5106065181930338 'Medium']
 [1.0 0.0 0.0 … 1.8354629319820521 -0.7959345330293706 'Low']
 [1.0 0.0 0.0 … 1.4143217110983428 -0.7959345330293706 'Low']
 …
 [1.0 0.0 0.0 … -0.6792731303453047 -0.6418293606808149 'Medium']
 [1.0 0.0 0.0 … 1.053742394902809 -0.7959345330293706 'Low']
 [0.0 1.0 0.0 … 1.261517632686496 -0.7959345330293706 'Low']]
```

[10]:
```
# prepare data frame for splitting data into train and test datasets

features = []
features = df_churn_pd.drop(['CHURNRISK'], axis=1)

label_churn = pd.DataFrame(df_churn_pd, columns = ['CHURNRISK'])
label_encoder = LabelEncoder()
label = df_churn_pd['CHURNRISK']

label = label_encoder.fit_transform(label)
print("Encoded value of Churnrisk after applying label encoder : " + str(label))
```

```
Encoded value of Churnrisk after applying label encoder : [2 1 1 … 2 1 1]
```

[11]:
```
area = 75
x = df_churn_pd['ESTINCOME']
y = df_churn_pd['DAYSSINCELASTTRADE']
z = df_churn_pd['TOTALDOLLARVALUETRADED']

pop_a = mpatches.Patch(color='#BB6B5A', label='High')
pop_b = mpatches.Patch(color='#E5E88B', label='Medium')
pop_c = mpatches.Patch(color='#8CCB9B', label='Low')
def colormap(risk_list):
    cols=[]
    for l in risk_list:
        if l==0:
            cols.append('#BB6B5A')
        elif l==2:
            cols.append('#E5E88B')
        elif l==1:
            cols.append('#8CCB9B')
    return cols

fig = plt.figure(figsize=(12,6))
fig.suptitle('2D and 3D view of churnrisk data')

# First subplot
ax = fig.add_subplot(1, 2,1)

ax.scatter(x, y, alpha=0.8, c=colormap(label), s= area)
ax.set_ylabel('DAYS SINCE LAST TRADE')
ax.set_xlabel('ESTIMATED INCOME')

plt.legend(handles=[pop_a,pop_b,pop_c])

# Second subplot
ax = fig.add_subplot(1,2,2, projection='3d')
```

```
ax.scatter(z, x, y, c=colormap(label), marker='o')

ax.set_xlabel('TOTAL DOLLAR VALUE TRADED')
ax.set_ylabel('ESTIMATED INCOME')
ax.set_zlabel('DAYS SINCE LAST TRADE')

plt.legend(handles=[pop_a,pop_b,pop_c])

plt.show()
```



[12]:
```
X_train, X_test, y_train, y_test = train_test_split(features, label,␣
↪random_state=0)
print("Dimensions of datasets that will be used for training : Input features"␣
↪+ str(X_train.shape) + " Output label" + str(y_train.shape))
print("Dimensions of datasets that will be used for testing : Input features" +␣
↪str(X_test.shape) + " Output label" + str(y_test.shape))
```

```
Dimensions of datasets that will be used for training : Input features(1549, 15)
Output label(1549,)
Dimensions of datasets that will be used for testing : Input features(517, 15)
Output label(517,)
```

[13]:
```
model_name = "Random Forest Classifier"

randomForestClassifier = RandomForestClassifier(n_estimators=100, max_depth=2,␣
↪random_state=0)
```

[14]:
```
rfc_model = Pipeline(steps=[('preprocessorAll', preprocessorForAllColumns),␣
 ↪('classifier', randomForestClassifier)])
```

[15]:
```
# Build models

rfc_model.fit(X_train, y_train)
```

[15]:
```
Pipeline(steps=[('preprocessorAll',
                 ColumnTransformer(remainder='passthrough',
                                   transformers=[('cat',
                                                  Pipeline(steps=[('impute',
SimpleImputer(strategy='most_frequent')),
                                                                  ('onehot',
OneHotEncoder(handle_unknown='ignore'))]),
                                                  ['GENDER', 'STATUS',
                                                   'HOMEOWNER']),
                                                 ('num',
                                                  Pipeline(steps=[('scale',
StandardScaler())]),
                                                  Index(['CHILDREN',
'ESTINCOME', 'AGE', 'TOTALDOLLARVALUETRADED',
       'TOTALUNITSTRADED', 'LARGESTSINGLETRANSACTION',
       'SMALLESTSINGLETRANSACTION', 'PERCENTCHANGECALCULATION',
       'DAYSSINCELASTLOGIN', 'DAYSSINCELASTTRADE', 'NETREALIZEDGAINS_YTD',
       'NETREALIZEDLOSSES_YTD'],
      dtype='object'))])),
                ('classifier',
                 RandomForestClassifier(max_depth=2, random_state=0))])
```

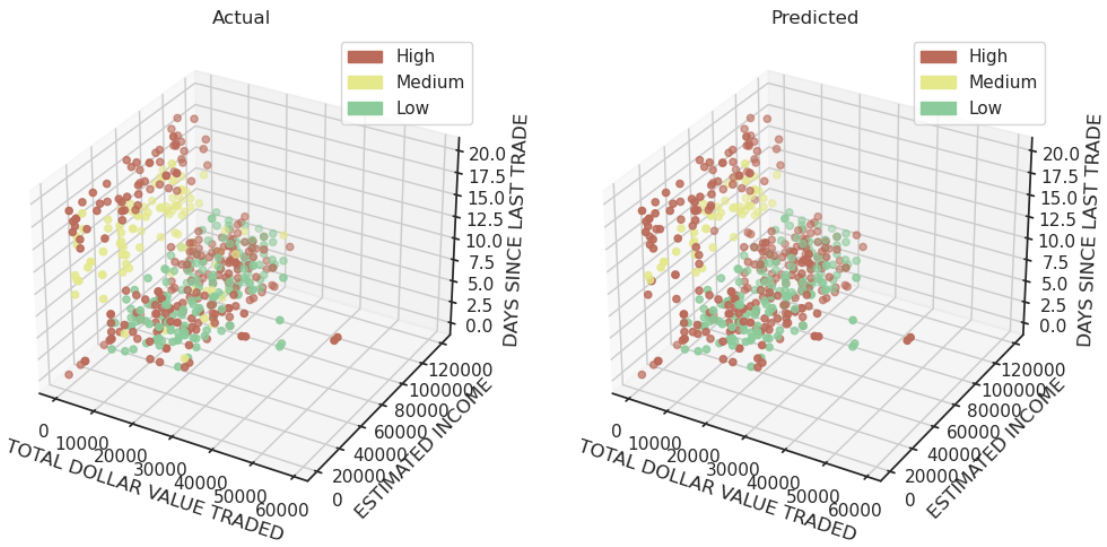
[16]:
```
y_pred_rfc = rfc_model.predict(X_test)
```

[17]:
```
def colormap(risk_list):
    cols=[]
    for l in risk_list:
        if l==0:
            cols.append('#BB6B5A')
        elif l==2:
            cols.append('#E5E88B')
        elif l==1:
            cols.append('#8CCB9B')
    return cols

def two_d_compare(y_test,y_pred,model_name):
    #y_pred = label_encoder.fit_transform(y_pred)
    #y_test = label_encoder.fit_transform(y_test)
    area = (12 * np.random.rand(40))**2
    plt.subplots(ncols=2, figsize=(10,4))
```

```
    plt.suptitle('Actual vs Predicted data : ' +model_name + '. Accuracy : %.
 ↪2f' % accuracy_score(y_test, y_pred))

    plt.subplot(121)
    plt.scatter(X_test['ESTINCOME'], X_test['DAYSSINCELASTTRADE'], alpha=0.8,␣
 ↪c=colormap(y_test))
    plt.title('Actual')
    plt.legend(handles=[pop_a,pop_b,pop_c])

    plt.subplot(122)
    plt.scatter(X_test['ESTINCOME'], X_test['DAYSSINCELASTTRADE'],alpha=0.8,␣
 ↪c=colormap(y_pred))
    plt.title('Predicted')
    plt.legend(handles=[pop_a,pop_b,pop_c])

    plt.show()

x = X_test['TOTALDOLLARVALUETRADED']
y = X_test['ESTINCOME']
z = X_test['DAYSSINCELASTTRADE']

pop_a = mpatches.Patch(color='#BB6B5A', label='High')
pop_b = mpatches.Patch(color='#E5E88B', label='Medium')
pop_c = mpatches.Patch(color='#8CCB9B', label='Low')

def three_d_compare(y_test,y_pred,model_name):
    fig = plt.figure(figsize=(12,10))
    fig.suptitle('Actual vs Predicted (3D) data : ' +model_name + '. Accuracy :␣
 ↪%.2f' % accuracy_score(y_test, y_pred))

    ax = fig.add_subplot(121, projection='3d')
    ax.scatter(x, y, z, c=colormap(y_test), marker='o')
    ax.set_xlabel('TOTAL DOLLAR VALUE TRADED')
    ax.set_ylabel('ESTIMATED INCOME')
    ax.set_zlabel('DAYS SINCE LAST TRADE')
    plt.legend(handles=[pop_a,pop_b,pop_c])
    plt.title('Actual')

    ax = fig.add_subplot(122, projection='3d')
    ax.scatter(x, y, z, c=colormap(y_pred), marker='o')
    ax.set_xlabel('TOTAL DOLLAR VALUE TRADED')
    ax.set_ylabel('ESTIMATED INCOME')
    ax.set_zlabel('DAYS SINCE LAST TRADE')
    plt.legend(handles=[pop_a,pop_b,pop_c])
    plt.title('Predicted')

    plt.show()
```

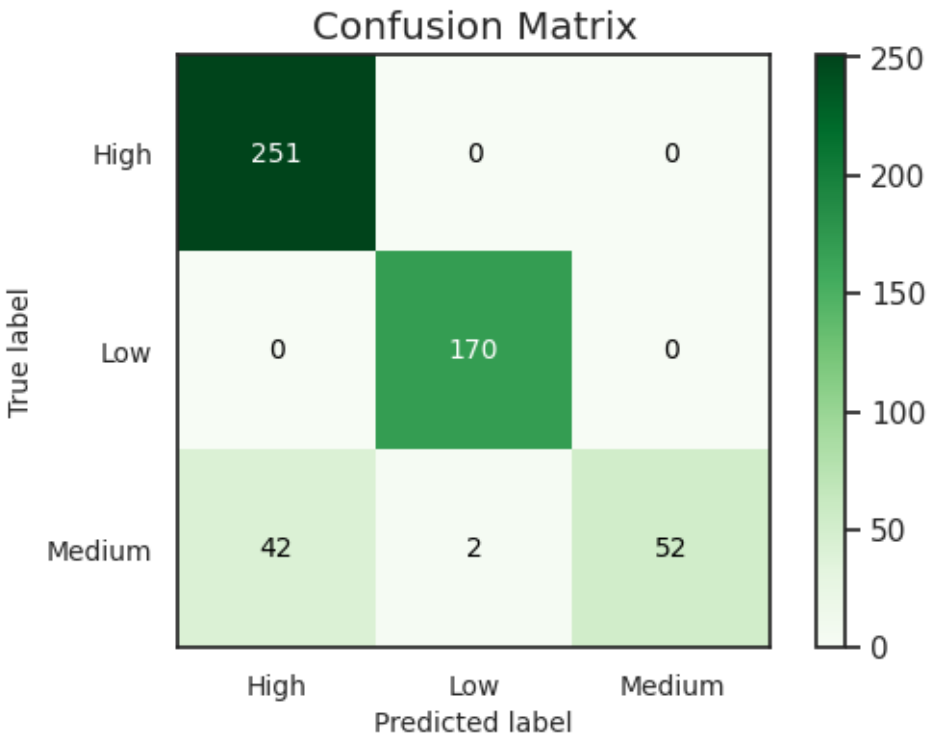
```
def model_metrics(y_test,y_pred):
    print("Decoded values of Churnrisk after applying inverse of label encoder :
 ↪ " + str(np.unique(y_pred)))

    skplt.metrics.
 ↪plot_confusion_matrix(y_test,y_pred,text_fontsize="small",cmap='Greens',figsize=(6,4))
    plt.show()

    print("The classification report for the model : \n\n"+␣
 ↪classification_report(y_test, y_pred))
```

[18]:
```
two_d_compare(y_test, y_pred_rfc, model_name)

three_d_compare(y_test,y_pred_rfc,model_name)
```



[19]:
```
y_test = label_encoder.inverse_transform(y_test)
y_pred_rfc = label_encoder.inverse_transform(y_pred_rfc)
model_metrics(y_test, y_pred_rfc)
```

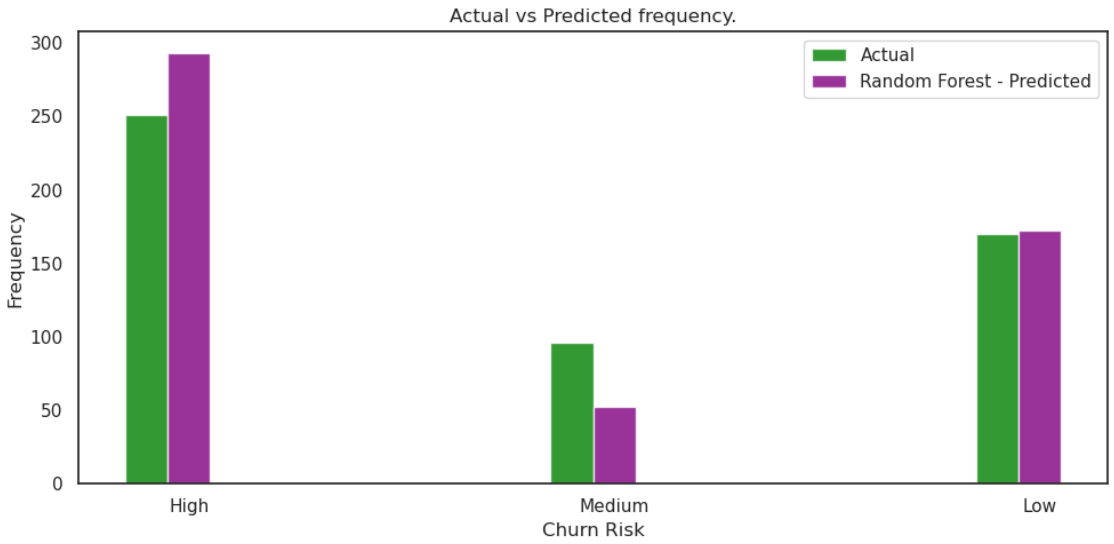
```
Decoded values of Churnrisk after applying inverse of label encoder : ['High'
'Low' 'Medium']
```




```
The classification report for the model :

              precision    recall  f1-score   support

        High       0.86      1.00      0.92       251
         Low       0.99      1.00      0.99       170
      Medium       1.00      0.54      0.70        96

    accuracy                           0.91       517
   macro avg       0.95      0.85      0.87       517
weighted avg       0.93      0.91      0.91       517
```

[20]:
```
uniqueValues, occurCount = np.unique(y_test, return_counts=True)
frequency_actual = (occurCount[0],occurCount[2],occurCount[1])

uniqueValues, occurCount = np.unique(y_pred_rfc, return_counts=True)
frequency_predicted_rfc = (occurCount[0],occurCount[2],occurCount[1])

n_groups = 3
fig, ax = plt.subplots(figsize=(10,5))
index = np.arange(n_groups)
bar_width = 0.1
```

```
opacity = 0.8

rects1 = plt.bar(index, frequency_actual, bar_width,
alpha=opacity,
color='g',
label='Actual')

rects6 = plt.bar(index + bar_width, frequency_predicted_rfc, bar_width,
alpha=opacity,
color='purple',
label='Random Forest - Predicted')

plt.xlabel('Churn Risk')
plt.ylabel('Frequency')
plt.title('Actual vs Predicted frequency.')
plt.xticks(index + bar_width, ('High', 'Medium', 'Low'))
plt.legend()

plt.tight_layout()
plt.show()
```



[ ]: