

Data Intake Report Insurance

Name: Insurance

Report date: 25 January 2022

Internship Batch: LISUM05

Version: 1.0

Data intake by: Selaelo Kholofelo Ramokgopa

Data intake reviewer:

Data storage location: computer

Tabular data details:

Total number of observations	1338
Total number of files	1
Total number of features	7
Base format of the file	.CSV
Size of the data	73.3 KB

Proposed Approach: Data pre-processing

- Information on the dataset was obtained
- Dataset was checked for null values and duplicates
- Column headers were viewed
- Pair wise relationships in the data were plotted
- Flask deployment
 - Xgboost model was selected and fit to data

```
1250    18648.42170
1257    11305.93455
139      2166.73200
919      5245.22690
...
298      38746.35510
954      20009.63365
950      11534.87265
399      1631.66830
1217     4058.11610
Name: charges, Length: 402, dtype: float64
```

```
In [21]: model.fit(x, y)
```

```
[05:52:46] WARNING: ..\src\learner.cc:1115: Starting in XGBoost 1.3.0, the default evaluation metric used with the objective 'multi:softprob' was changed from 'merror' to 'mlogloss'. Explicitly set eval_metric if you'd like to restore the old behavior.
```

```
Out[21]: XGBClassifier(base_score=0.5, booster='gbtree', colsample_bylevel=1,
                      colsample_bynode=1, colsample_bytree=1, enable_categorical=False,
                      gamma=0, gpu_id=-1, importance_type=None,
                      interaction_constraints='', learning_rate=0.300000012,
                      max_delta_step=0, max_depth=6, min_child_weight=1, missing=nan,
                      monotone_constraints='()', n_estimators=100, n_jobs=4, nthread=-1,
                      num_parallel_tree=1, objective='multi:softprob', predictor='auto',
                      random_state=1, reg_alpha=0, reg_lambda=1, scale_pos_weight=None,
                      seed=1, subsample=1, tree_method='exact', validate_parameters=1, ...)
```

```
In [22]: # Save the model
```

-
- App.py file was created

```
File Edit Search Source Run Debug Consoles Projects Tools View Help

C:\Users\SelaLo1\app.py

temp.py x app.py x kite_tutorial.py x index.html x

1  # -*- coding: utf-8 -*-
2  """
3  Created on Wed Jan 19 23:26:27 2022
4
5  @author: SelaLo1
6  """
7
8  from flask import Flask, request, url_for, redirect, render_template, jsonify
9  import xgboost as xgb
10 import pandas as pd
11 import pickle
12 import numpy as np
13
14 # Initialise the Flask app
15 app = Flask(__name__)
16
17 # Loads pre-trained model
18 model = pickle.load(open('model.sav', 'rb'))
19
20 cols = ['age', 'sex', 'bmi', 'children', 'smoker', 'region']
21
22 @app.route('/')
23 def home():
24     return render_template("index.html")
25
26 @app.route('/predict', methods=['POST'])
27 def predict():
28     int_features = [x for x in request.form.values()]
29     final = np.array(int_features)
30     data_unseen = pd.DataFrame([final], columns = cols)
31     prediction = predict(model, data=data_unseen, round = 0)
32     prediction = int(prediction.Label(0))
33     return render_template("index.html", pred='Expected Bill will be {}'.format(pre
34
35 @app.route('/predict_api', methods=['POST'])
36 def predict_api():
37     data = request.get_json(force=True)
38     data_unseen = pd.DataFrame(data)
```

-
-

Html file was created

```
File Edit Search Source Run Debug Consoles Projects Tools View Help

C:\Users\SelaLo1\spyder-py3\autosave\index.html

temp.py x app.py x kite_tutorial.py x index.html x

1  <!DOCTYPE html>
2  <html lang="en">
3
4  <!-- From https://codepen.io/frtyler/pen/EGdtg-->
5
6  <title>ML API</title>
7  <link href="https://fonts.googleapis.com/css?family=Pacifico" rel="stylesheet" type="text/css">
8  <link href="https://fonts.googleapis.com/css?family=Acima" rel="stylesheet" type="text/css">
9  <link href="https://fonts.googleapis.com/css?family=Hind:300" rel="stylesheet" type="text/css">
10 <link href="https://fonts.googleapis.com/css?family=Open+Sans+Condensed:300" rel="stylesheet" type="text/css">
11 <link rel="stylesheet" href="{{ url_for('static', filename='css/style.css') }}">
12
13 <head>
14     <meta charset="UTF-8">
15     <meta name="viewport" content="width=device-width, initial-scale=1.0">
16     <title>Document</title>
17 </head>
18
19 <body>
20
21     <div style="color:blue">
22         <form action="{{ url_for('predict') }}" method="post">
23             <h2>Predict insurance price</h2>
24
25             <h3>What is your age</h3><br><input id="first" name="age" required="required" type="text">
26             <h3>What is your sex</h3><br><select id="second" name="sex" required="required">
27                 <option value="Male">Male</option>
28                 <option value="Female">Female</option>
29             </select>
30             <h3>What is your bmi</h3><br><input id="third" name="bmi" required="required" type="text">
31             <h3>How many children do you have</h3><br><input name="children" id="four" type="text">
32             <h3>Do you smoke</h3><br><select name="smoke" id="fifth" required="required">
33                 <option value="Yes">Yes</option>
34                 <option value="No">No</option>
35             </select>
36             <h3>Which region are you from</h3><br><select name="region" id="sixth" type="text">
37
38         </form>
```

-

Predict insurance price

What is your age

What is your sex

Male ▾

What is your bmi

How many children do you have

Do you smoke

Yes ▾

Which region are you from

southwest ▾

-
-

Requirement.txt was created

```
requirements - Notepad
File Edit Format View Help
alabaster==0.7.12
anaconda-client==1.7.2
anaconda-navigator==1.9.12
anaconda-project==0.8.3
argh==0.26.2
asn1crypto==1.3.0
astroid==2.3.3
astropy==4.0
atomicwrites==1.3.0
attrs==19.3.0
autopep8==1.4.4
Babel==2.8.0
backcall==0.1.0
backports.functools-lru-cache==1.6.4
backports.shutil-get-terminal-size==1.0.0
backports.tempfile==1.0
backports.weakref==1.0.post1
bcrypt==3.1.7
beautifulsoup4==4.8.2
bitarray==1.2.1
bkcharts==0.2
bleach==3.1.0
bokeh==1.4.0
boto==2.49.0
Bottleneck==1.3.2
certifi==2019.11.28
cffi==1.14.0
chardet==3.0.4
charset-normalizer==2.0.10
Click==7.0
cloudpickle==1.3.0
clyent==1.2.2
colorama==0.4.3
colorlover==0.3.0
comtypes==1.1.7
conda==4.11.0
conda-build==3.18.11
conda-package-handling==1.7.3
<
```

-
-

Model was deployed in flask

```
Desktop  Concession  Downloads  Khanyi Thesis  UC Emblem
Anaconda Prompt (anaconda3) - python app.py

(base) C:\Users\Selaelo1> cd Environments

(base) C:\Users\Selaelo1\Environments>cd my_env

(base) C:\Users\Selaelo1\Environments\my_env> python app.py
* Serving Flask app "app" (lazy loading)
* Environment: production
  WARNING: This is a development server. Do not use it in a production deployment.
  Use a production WSGI server instead.
* Debug mode: on
* Restarting with windowsapi reloader
* Debugger is active!
* Debugger PIN: 179-618-667
* Running on http://127.0.0.1:5000/ (Press CTRL+C to quit)
```