SoFi

Fwd: The Gen3 score variable on MBE

Hua Peng <hpeng@sofi.org>   Wed, Mar 10, 2021 at 9:50 AM
To: George Xu <jxu@sofi.org>

Hua Peng

Staff Data Scientist, Modeler

234 First Street

San Francisco, CA 94129

Email: hpeng@sofi.org

SoFi

---------- Forwarded message ---------
From: **Bo Shao** <bshao@sofi.org>
Date: Tue, Mar 9, 2021 at 10:41 PM
Subject: Fwd: The Gen3 score variable on MBE
To: Hua Peng <hpeng@sofi.org>

Bo Shao
Senior Data Scientist, Risk Data Science
San Francisco, CA

---------- Forwarded message ---------
From: **Bo Shao** <bshao@sofi.org>
Date: Fri, Nov 6, 2020 at 1:11 PM
Subject: Re: The Gen3 score variable on MBE
To: Jessica Liao <jessica.liao@creditkarma.com>
Cc: Luke Zhang <lzhang@sofi.org>

Finally, it seems to be working and passed the validation.
for one last step, we need to set score the people with missing attributes as 'invalid score' or "-88"
Please let us know when you finished the pre_processing csv file.

Thanks,
Bo Shao
Senior Data Scientist, Risk Data Science
San Francisco, CA

On Fri, Nov 6, 2020 at 11:29 AM Jessica Liao <jessica.liao@creditkarma.com> wrote:

> On Fri, Nov 6, 2020 at 10:58 AM Bo Shao <bshao@sofi.org> wrote:
>> Jessica,
>> I tried formula, but failed. Maybe you can show us how to do it during the call.
>>
>> LOCAL   ROUND(LN((probability(0.0))/probability(1.0))*132.877+557.123,0)   gen3_score
>>
>> Thanks,
>> Bo Shao
>> Senior Data Scientist, Risk Data Science
>> San Francisco, CA
>>
>> On Fri, Nov 6, 2020 at 10:42 AM Luke Zhang <lzhang@sofi.org> wrote:
>>> Works for me. Let's meet at 11 then. Please find the link below to join.
>>>
>>> https://sofi.zoom.us/j/9148338830

On Fri, Nov 6, 2020 at 10:39 AM Bo Shao <bshao@sofi.org> wrote:
> I am good at 11:00
>
> Bo Shao
> Senior Data Scientist, Risk Data Science
> San Francisco, CA

On Fri, Nov 6, 2020 at 10:37 AM Jessica Liao <jessica.liao@creditkarma.com> wrote:
> Wow, totally jealous of that!! Any 30 minutes between 11:00AM-12:00PM also works for me. Let me know and I will send out invites!
>
> Thanks!
> Jessica

On Fri, Nov 6, 2020 at 10:29 AM Bo Shao <bshao@sofi.org> wrote:
> Do you have time to meet before 2pm? SoFi is off after 2pm Friday.
>
> Thanks,
> Bo Shao
> Senior Data Scientist, Risk Data Science
> San Francisco, CA

On Fri, Nov 6, 2020 at 10:25 AM Jessica Liao <jessica.liao@creditkarma.com> wrote:
> Hi Bo, Luke,
>
> I now have the model scoring and outputting the gen3 score to be used in post csv with no issues, do you have some time later today to meet the catch up? We'll need to make changes in both pre-csv and the pmml. Please see attached files I used for testing purposes. One call out is the pre_csv is set as no missing imputation for all inputs, which need to be modified - We can talk it through over the call.
>
> I am available from 3PM-5PM today. 30 minutes should work!! Let me know when it works best for you!! Talk soon!
>
> Thanks!
> Jessica

On Thu, Nov 5, 2020 at 9:46 PM Bo Shao <bshao@sofi.org> wrote:
> Tried a lot of ways , but all failed.
> How about directly editing the old-version XML file(which works)? Can MBE take the output from either 'probability_1' or 'Predicited_y' or any given name?
> or any idea about how to edit this output part to make it work?

```
                </Segmentation>
            </MiningModel>
        <Segment id="2">
            <True/>
            <RegressionModel functionName="classification" normalizationMethod="logit">
                <MiningSchema>
                    <MiningField name="lgbmValue"/>
                </MiningSchema>
                <Output>
                    <OutputField name="probability_0" optype="continuous" dataType="double" feature="probability" value="0.0"/>
                    <OutputField name="probability_1" optype="continuous" dataType="double" feature="probability" value="1.0"/>
                    <OutputField name="Predicted_y" optype="continuous" dataType="double" feature="transformedValue">
                        <FieldRef field="probability_1"/>
                    </OutputField>
                </Output>
                <RegressionTable intercept="0.0" targetCategory="1.0">
                    <NumericPredictor name="lgbmValue" coefficient="1.0"/>
                </RegressionTable>
                <RegressionTable intercept="0.0" targetCategory="0.0"/>
            </RegressionModel>
        </Segment>
    /Segmentation>
odel>
```

> Thanks,
> Bo Shao
> Senior Data Scientist, Risk Data Science
> San Francisco, CA

On Thu, Nov 5, 2020 at 8:33 PM Jessica Liao <jessica.liao@creditkarma.com> wrote:
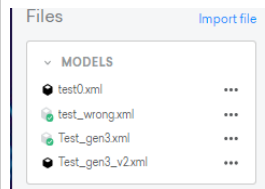> Hi Bo,
>
> All models show up in MBE MODELS section were compiled successfully, and the icon will only turn to a green check after you use that compiled model to score a df in MBE successfully! Do you remember what the manual edits you made in the previous version of XML file to make it scoring? Can you elaborate on that? Thanks!
>
> Thanks!
> Jessica

On Thu, Nov 5, 2020 at 7:48 PM Bo Shao <bshao@sofi.org> wrote:

Jessica,

After updating pmml version to 4.3 and trying compiling again, the scoring still fails. I remembered when I did successfully compiling the old version, I did some manuall editing in the XML file, Is this the same issue?
And i noticed that the icon of the new generated xml file is not green, does it mean it is not compiled successfully?

btw, I am able to use the pmml file to score in my environment by using

from pypmml import Model
model = Model.fromFile("gen3_pl_Lightbox_pmml.xml")
train_eval1['pred2']=model.predict(train_eval1[lllx])['probability(1.0)'].values

Thanks,
Bo Shao
Senior Data Scientist, Risk Data Science
San Francisco, CA

On Thu, Nov 5, 2020 at 5:53 PM Jessica Liao <jessica.liao@creditkarma.com> wrote:

Hi Bo,

The error you got is because of the PMML version. MBE currently only supports for PMML version 4.3 while your PMML version is 4.4. You can open the pmml file and check the version at the top few lines. The sklearn2pmml package uses the "JPMML-SkLearn" library for actual conversion work. And your "JPMML-SkLearn" version is "1.6.7" (newest version) So you may need to Uninstall that and install an older version, eg version 1.5.x

Let me know if that works!

Thanks!
Jessica

On Thu, Nov 5, 2020 at 1:52 PM Bo Shao <bshao@sofi.org> wrote:

Jessica,

I am able to generate PMML file and finished compile in MBE, but it cannot do scoring,
Please see the errors below. Also, I attached PMML file and transformation file.   And if you can access out MBE, please see
""Model testing_sample_test".

```
import cklearn.v1 as ck
ck.compile(files=["A_data_transformations_v6.csv","gen3_pl_Lightbox_pmml.xml"], model_name="Test_gen3_v2.xml")

  Info: Compile: Start ---------------
  Info: A_data_transformations_v6.csv: Processing
  Info: gen3_pl_Lightbox_pmml.xml: Processing
  Info: Chaining PMML
  Info: Compile: End ----------------
  Info: Output file Test_gen3_v2.xml generated
```

```
[0]: my_df_prescored = ck.csv_to_dataframe(filename="test_12.csv")
     my_df_scored = ck.score_model(model_name="Test_gen3_v2.xml", df=my_df_prescored)

    Converting to dataframe.
    Elapsed: 0m 1s.  Memory usage: 4%.  Checking again in 2 seconds...
    Finished.  Total time: 0m 3s.  Memory usage: 4%
    Start scoring 10048 rows 2020-11-05 20:16:05.020034
    Starting scoring for rows  0 to  10047 2020-11-05 20:16:05.211481

    ---------------------------------------------------------------------
    _RemoteTraceback                        Traceback (most recent call last)
    _RemoteTraceback:
    """
    Traceback (most recent call last):
      File "/usr/local/envs/py3env/lib/python3.5/site-packages/loky/process_executor.py", line 431, in _process_worker
        r = call_item()
      File "/usr/local/envs/py3env/lib/python3.5/site-packages/loky/process_executor.py", line 285, in __call__
        return self.fn(*self.args, **self.kwargs)
      File "/usr/local/envs/py3env/lib/python3.5/site-packages/cklearn/pmml_utilities/__init__.py", line 365, in evaluate_csv_jpmml
        shell._execute_bash_command(jpmml_cmd)
      File "/usr/local/envs/py3env/lib/python3.5/site-packages/cklearn/shell_utilities/__init__.py", line 28, in __execute_bash_command
        universal_newlines=True
      File "/usr/local/envs/py3env/lib/python3.5/subprocess.py", line 316, in check_output
        **kwargs).stdout
      File "/usr/local/envs/py3env/lib/python3.5/subprocess.py", line 398, in run
        output=stdout, stderr=stderr)
    subprocess.CalledProcessError: Command '['java', '-Xmx100G', '-cp', '/jpmml-evaluator/pmml-evaluator-example/target/example-1.4.1.jar', 'org.jpmml.
    evaluator.EvaluationExample', '--model', 'tmp_71431a62-284c-47ea-857d-6980e460ef6d/model.pmml', '--input', 'tmp_71431a62-284c-47ea-857d-6980e460ef6
    d/inputs/aa', '--output', 'tmp_71431a62-284c-47ea-857d-6980e460ef6d/outputs/aa']' returned non-zero exit status 1
    """

    The above exception was the direct cause of the following exception:

    CalledProcessError                      Traceback (most recent call last)
    <ipython-input-30-40c8944b11c9> in <module>()
          1 my_df_prescored = ck.csv_to_dataframe(filename="test_12.csv")
    ----> 2 my_df_scored = ck.score_model(model_name="Test_gen3_v2.xml", df=my_df_prescored)

    /usr/local/envs/py3env/lib/python3.5/site-packages/cklearn/metric_logging/__init__.py in wrapped_func(*args, **kwargs)
         65         result = _SUCCESS
         66         try:
    ---> 67             return func(*args, **kwargs)
         68         except Exception as e:
         69             result = e.__class__.__name__

    /usr/local/envs/py3env/lib/python3.5/site-packages/cklearn/v1/__init__.py in score_model(model_name, df, sparse, workspace_id, **kwargs)
        217             model_name=model_name,
        218             scoring_input=scoring_input,
    --> 219             sparse=sparse, **kwargs)
        220
        221

    /usr/local/envs/py3env/lib/python3.5/site-packages/cklearn/model_validation/scoring.py in do_score_model(workspace_id, model_name, scoring_input, **kwargs)
         45             scoring_data,
         46             update_ck_validation_bool=True,
    ---> 47             **kwargs)
         48
         49

    /usr/local/envs/py3env/lib/python3.5/site-packages/cklearn/model_validation/scoring.py in __score_model_local(workspace_id, model_name, pmml, input_csv, update_ck_validation_bool, **kwargs)
         77         score_distribution_plot = kwargs.pop('score_distribution_plot', True)
         78
    ---> 79         ret_df = evaluate_csv(pmml=pmml, input_csv=input_csv, **kwargs)
         80
         81         if score_distribution_plot:

    /usr/local/envs/py3env/lib/python3.5/site-packages/cklearn/pmml_utilities/__init__.py in evaluate_csv(pmml, input_csv, sparse, chunk_size, use_parallel_scoring)
        149
        150         score_inputs_and_save_result(chunk_size, inputs_dir, model_filename, outputs_dir, sparse,
    --> 151                                      use_parallel_scoring = use_parallel_scoring)
        152
        153         # Combine output files into one dataframe

    /usr/local/envs/py3env/lib/python3.5/site-packages/cklearn/pmml_utilities/__init__.py in score_inputs_and_save_result(num_records_per_input_file, input_files_dir, model_filename, outputs_dir, sparse, use_parallel_scoring)
        185         score_inputs_in_parallel_and_save_result(num_records_per_input_file = num_records_per_input_file,
        186                                      input_files_dir = input_files_dir, model_filename = model_filename,
    --> 187                                      outputs_dir = outputs_dir, sparse = sparse)
        188     else:
        189         score_inputs_sequentially_and_save_result(num_records_per_input_file = num_records_per_input_file,

    /usr/local/envs/py3env/lib/python3.5/site-packages/cklearn/pmml_utilities/__init__.py in score_inputs_in_parallel_and_save_result(num_records_per_input_file, input_files_dir, model_filename, outputs_dir, sparse)
        254         # of actually propagating it upstream
        255         if thrown_exceptions:
    --> 256             raise thrown_exceptions.pop()
        257
        258

    CalledProcessError: Command '['java', '-Xmx100G', '-cp', '/jpmml-evaluator/pmml-evaluator-example/target/example-1.4.1.jar', 'org.jpmml.evaluator.EvaluationExample', '--model', 'tmp_71431a62-284c-47e
    a-857d-6980e460ef6d/model.pmml', '--input', 'tmp_71431a62-284c-47ea-857d-6980e460ef6d/inputs/aa', '--output', 'tmp_71431a62-284c-47ea-857d-6980e460ef6d/outputs/aa']' returned non-zero exit status 1
```

Thanks,
Bo Shao
Senior Data Scientist, Risk Data Science
San Francisco, CA

On Wed, Nov 4, 2020 at 10:52 PM Jessica Liao <jessica.liao@creditkarma.com> wrote:

Hi Bo,

The error message seems to point to the Java version, you may need to Uninstall the current java.exe, and install the latest one. Besides that, you can also try set with_repr and debug parameter as true when running sklearn2pmml as below:

```
sklearn2pmml(pipeline, 'gen3_pl_Lightbox_pmml.xml', with_repr=True, debug=True)
```

Let me know if this would help resolve the error you saw. Thanks!

Thanks!
Jessica

On Wed, Nov 4, 2020 at 10:05 PM Bo Shao <bshao@sofi.org> wrote:

Error message again

```
dtype: object
```

```
[8]:  from sklearn2pmml.decoration import Alias
      from  sklearn2pmml.preprocessing import ExpressionTransformer

      classifier = LGBMClassifier(random_state = 13)
      classifier = LGBMClassifier(objective = "binary",min_child_samples =300,boosting_type='gbdt',
                                  random_state=157,learning_rate=0.05,monotone_constraints=mono_list_constr,
                                  max_depth=6, n_estimators = 800)

      pipeline = PMMLPipeline([
        ("mapper",mapper),
        ("classifier", classifier)],
        predict_proba_transformer = Alias(ExpressionTransformer("X[1]"), name="Predicted_y", prefit=True))
      pipeline.fit(df_X, df_y, classifier__sample_weight=df_w)
```

```
[8]:  PMMLPipeline(steps=[('mapper', DataFrameMapper(drop_cols=[],
                           features=[(['ALL7938'],
                                      ContinuousDomain(high_value=100.0, low_value=0.0,
                                                       outlier_treatment='as_missing_values')),
                                     (['TBCC3351T'],
                                      ContinuousDomain(high_value=990.0, low_value=0.0,
                                                       outlier_treatment='as_missing_values')),
                                     (['TBCC2502T'],
                                      ContinuousDomain(high_value=9990.0, low_value=0.0,
                                                       outlier_treatment='as_missing_values')),
                                     (['...
                                                       outlier_treatment='as_missing_values')),
                                     (['TBCA2618T'],
                                      ContinuousDomain(high_value=90.0, low_value=0.0,
                                                       outlier_treatment='as_missing_values')),
                                     (['FIP5020'],
                                      ContinuousDomain(high_value=999999990.0,
                                                       low_value=0.0,
                                                       outlier_treatment='as_missing_values')),
                                     (['PIL0438'],
                                      ContinuousDomain(high_value=90.0, low_value=0.0,
                                                       outlier_treatment='as_missing_values')), ...])),
                    ('classifier', LGBMClassifier(learning_rate=0.05, max_depth=6, min_child_samples=300,
                           monotone_constraints=[-1.0, 1.0, 1.0, -1.0, -1.0, 1.0, -1.0, 1.0,
                                                 1.0, -1.0, 1.0, -1.0, -1.0, 1.0, 1.0, 1.0,
                                                 1.0, -1.0, -1.0, -1.0, 1.0, 1.0, -1.0, -1.0,
                                                 -1.0, 1.0, -1.0, 1.0, 1.0, 1.0, 1.0, ...],
                           n_estimators=800, objective='binary', random_state=157))])
```

```
[ ]:  from sklearn2pmml import sklearn2pmml
      sklearn2pmml(pipeline, "gen3_pl_lightbox_pmml.xml")

      Standard output is empty
      Standard error:
      Exception in thread "main" java.lang.UnsupportedClassVersionError: org/jpmml/sklearn/Main : Unsupported major.minor version 52.0
              at java.lang.ClassLoader.defineClass1(Native Method)
              at java.lang.ClassLoader.defineClass(ClassLoader.java:808)
              at java.security.SecureClassLoader.defineClass(SecureClassLoader.java:142)
              at java.net.URLClassLoader.defineClass(URLClassLoader.java:443)
              at java.net.URLClassLoader.access$100(URLClassLoader.java:65)
              at java.net.URLClassLoader$1.run(URLClassLoader.java:355)
              at java.net.URLClassLoader$1.run(URLClassLoader.java:349)
              at java.security.AccessController.doPrivileged(Native Method)
              at java.net.URLClassLoader.findClass(URLClassLoader.java:348)
              at java.lang.ClassLoader.loadClass(ClassLoader.java:430)
              at sun.misc.Launcher$AppClassLoader.loadClass(Launcher.java:326)
              at java.lang.ClassLoader.loadClass(ClassLoader.java:363)
              at sun.launcher.LauncherHelper.checkAndLoadMain(LauncherHelper.java:482)
```

```
---------------------------------------------------------------------------
RuntimeError                              Traceback (most recent call last)
<ipython-input-59-3a3e6e8188af> in <module>
      1 from sklearn2pmml import sklearn2pmml
----> 2 sklearn2pmml(pipeline, "gen3_pl_lightbox_pmml.xml")

~/.local/lib/python3.6/site-packages/sklearn2pmml/__init__.py in sklearn2pmml(pipeline, pmml, user_classpath, with_repr, debug, java_encoding)
    256                 print("Standard error is empty")
    257         if retcode:
--> 258             raise RuntimeError("The JPMML-SkLearn conversion application has failed. The Java executable should have printed more information about the failure into its standard out
put and/or standard error streams")
    259         finally:
    260             if debug:

RuntimeError: The JPMML-SkLearn conversion application has failed. The Java executable should have printed more information about the failure into its standard output and/or standard error streams
```

Bo Shao
Senior Data Scientist, Risk Data Science
San Francisco, CA

On Wed, Nov 4, 2020 at 8:42 PM Jessica Liao <jessica.liao@creditkarma.com> wrote:

Hi Bo,

You'll need to import *Alias* from *sklearn2pmml* package first as below:

from sklearn2pmml.decoration import Alias

Let me know if this resolves the issue. Thanks!

Thanks!
Jessica

On Wed, Nov 4, 2020 at 8:23 PM Bo Shao <bshao@sofi.org> wrote:

Sure.

```
]: classifier = LGBMClassifier(random_state = 13)
   classifier = LGBMClassifier(objective = "binary",min_child_samples =300,boosting_type='gbdt',
                               random_state=157,learning_rate=0.05,monotone_constraints=mono_list_constr,
                               max_depth=6, n_estimators = 800)

   pipeline = PMMLPipeline([
     ("mapper",mapper),
     ("classifier", classifier)],
     predict_proba_transformer = Alias(ExpressionTransformer("X[1]"), name="Predicted_y", prefit=True))
   pipeline.fit(df_X, df_y, classifier__sample_weight=df_w)
   -----------------------------------------------------------------------
   NameError                                 Traceback (most recent call last)
   <ipython-input-89-b2a04a7b2976> in <module>
         7   ("mapper",mapper),
         8   ("classifier", classifier)],
   ----> 9   predict_proba_transformer = Alias(ExpressionTransformer("X[1]"), name="Predicted_y", prefit=True))
        10 pipeline.fit(df_X, df_y, classifier__sample_weight=df_w)

   NameError: name 'Alias' is not defined
```

Bo Shao
Senior Data Scientist, Risk Data Science
San Francisco, CA

On Wed, Nov 4, 2020 at 2:57 PM Jessica Liao <jessica.liao@creditkarma.com> wrote:
> Can you share the screenshot of the codes that generate this error for us to troubleshoot? Thanks!!
>
> Thanks!
> Jessica
>
> On Wed, Nov 4, 2020 at 2:32 PM Bo Shao <bshao@sofi.org> wrote:
>> Hi Jessica,
>>
>> Error messages shows :
>> name 'Alias' is not defined
>>
>> Thanks,
>> Bo Shao
>> Senior Data Scientist, Risk Data Science
>> San Francisco, CA
>>
>>
>> On Wed, Nov 4, 2020 at 12:40 PM Luke Zhang <lzhang@sofi.org> wrote:
>>> Thanks Jessica! Bo will do the step 1 and I will do step 2 and 3. It may take some time for Bo to do it
>>> because the person who originally did this left the company, but we will let you know when we check this
>>> out.
>>>
>>> Best,
>>> Luke
>>>
>>> On Wed, Nov 4, 2020 at 12:07 PM Jessica Liao <jessica.liao@creditkarma.com> wrote:
>>>> Hi Luke, Bo,
>>>>
>>>> Can you try below and see if it works:
>>>>
>>>> 1. Regenerate your risk model xml file by adding "predict_proba_transformer" in the PMMLPipeline as
>>>> following - Where X[1] will be the probability output and **you can name it whatever you want:**
>>>> PMMLPipeline([("mapper", mapper),
>>>>   ("classifier", classifier)],
>>>>   predict_proba_transformer = Alias(ExpressionTransformer("X[1]"), name="Predicted_y",
>>>> prefit=True))
>>>> 2. Compile a model with pre_csv + new pmml + post_rules.csv (attached) as a big model - let's call it
>>>> "test_model.xml"
>>>> 3. Score my_df_prescored with "test_model.xml" and check if the "X[1]" exists in the my_df_scored - you
>>>> can check all new added columns in the my_df_scored
>>>>
>>>> Let me know if you have any questions or concerns!
>>>>
>>>> Thanks!
>>>> Jessica
>>>>
>>>> On Tue, Nov 3, 2020 at 1:08 PM Luke Zhang <lzhang@sofi.org> wrote:
>>>>> Jessica, I aksed our DS team and please see attached the way they generate the PMML file.
>>>>>
>>>>> I included Bo in this email chain for any clarification questions on the model part.

```python
from lightgbm import LGBMClassifier
from sklearn2pmml.pipeline import PMMLPipeline
from sklearn2pmml.preprocessing.lightgbm import make_lightgbm_dataframe_mapper

df_X = train_main[lll]
df_y = train_main['target_bad']
df_w = train_main['weight']

mapper,_ = make_lightgbm_dataframe_mapper(df_X.dtypes, missing_value_aware = True)
classifier = LGBMClassifier(random_state = 13)
classifier = LGBMClassifier(objective = "binary",min_child_samples =300,boosting_type='gbdt',
                            random_state=157,learning_rate=0.05,monotone_constraints=mono_list_constr,
                            max_depth=6, n_estimators = 800)

pipeline = PMMLPipeline([
  ("mapper", mapper),
  ("classifier", classifier)
])
pipeline.fit(df_X, df_y, classifier__sample_weight=df_w)
```

```python
from sklearn2pmml import sklearn2pmml
sklearn2pmml(pipeline, "LightGBMtest_v1.xml")
```

On Tue, Nov 3, 2020 at 12:53 PM Jessica Liao <jessica.liao@creditkarma.com> wrote:
Thanks for confirming!! Can you also check with your DS team on what python package was used to generate the lightGBM model PMML file to help us better troubleshoot this issue? Thanks!!

Thanks!
Jessica

On Tue, Nov 3, 2020 at 12:46 PM Luke Zhang <lzhang@sofi.org> wrote:
Hi Jessica,

I tried the method above and found probability_1 not existed in the my_df_scored dataframe. Looks like it is still not outputted.

On Tue, Nov 3, 2020 at 12:36 PM Jessica Liao <jessica.liao@creditkarma.com> wrote:
Hi Luke,

Can you try following to confirm if the "probability_1" is outputted from the model pmml file:

- Replace your post csv file with an attached file when compiling the big model - let's call it "test_model.xml"
- Score my_df_prescored with "test_model.xml" and check if the "probability_1" exists in the my_df_scored dataframe

Thanks!
Jessica

On Tue, Nov 3, 2020 at 11:05 AM Jessica Liao <jessica.liao@creditkarma.com> wrote:

On Tue, Nov 3, 2020 at 10:45 AM Luke Zhang <lzhang@sofi.org> wrote:
Sounds good. Talk to you soon!

On Tue, Nov 3, 2020 at 10:43 AM Jessica Liao <jessica.liao@creditkarma.com> wrote:
Hi Luke,

Let's meet at 11:00AM, and I just sent out the invite. Talk soon!

Thanks!
Jessica

On Tue, Nov 3, 2020 at 10:37 AM Luke Zhang <lzhang@sofi.org> wrote:
That will be very helpful!! If we can convert the probability into a score in the csv, it will be the easiest way.

Let me know your availability today. I am free from now till 1:30 p.m, and from 3 p.m to 4 p.m.

Thanks,
Luke

On Tue, Nov 3, 2020 at 10:19 AM Jessica Liao <jessica.liao@creditkarma.com> wrote:
Hi Luke,

Happy Tuesday! What do you mean by adding another output (the score) to the model? Can you elaborate on that? Besides, do you need help to code the formula to generate the score from probability_1 in csv? Let me know if a working session today would be helpful! Thanks!! :)

Thanks!!
Jessica

On Tue, Nov 3, 2020 at 9:45 AM Luke Zhang <lzhang@sofi.org> wrote:
Hi Jessica,

I talked to our DS team and they said they only produced the probability_1 as the output but they have a formula to generate the score. Do you know how long it will take if they want to add another output (the score) to the model?

Thanks,

--
Luke Zhang
Senior Analyst, Risk Management
234 1st Street
San Francisco, CA 94105

SoFi

The information contained in this email message is PRIVATE and intended only for the personal and confidential use of the recipient named above. If the reader of this message is not the intended recipient or an agent responsible for delivering it to the intended recipient, you are hereby notified that you have received this message in error and that any review, dissemination, distribution or copying of this message is strictly prohibited. If you have received this communication in error, please notify us immediately by email, and delete the original message.

--
Jessica Liao | Partner Decision Scientist I | Credit Karma, Inc.
jessica.liao@creditkarma.com I www.creditkarma.com

This email may contain confidential and privileged information. Any review, use, distribution, or disclosure by anyone other than the intended recipient(s) is prohibited. If you are not the intended recipient, please contact the sender by reply email and delete all copies of this message.

--
Luke Zhang
Senior Analyst, Risk Management
234 1st Street
San Francisco, CA 94105

SoFi

The information contained in this email message is PRIVATE and intended only for the personal and confidential use of the recipient named above. If the reader of this message is not the intended recipient or an agent responsible for delivering it to the intended recipient, you are hereby notified that you have received this message in error and that any review, dissemination, distribution or copying of this message is strictly prohibited. If you have received this communication in error, please notify us immediately by email, and delete the original message.

--
Jessica Liao | Partner Decision Scientist I | Credit Karma, Inc.
jessica.liao@creditkarma.com I www.creditkarma.com

This email may contain confidential and privileged information. Any review, use, distribution, or disclosure by anyone other than the intended recipient(s) is prohibited. If you are not the intended recipient, please contact the sender by reply email and delete all copies of this message.

--

Luke Zhang
Senior Analyst, Risk Management
234 1st Street
San Francisco, CA 94105

SoFi

The information contained in this email message is PRIVATE and intended only for the personal and confidential use of the recipient named above. If the reader of this message is not the intended recipient or an agent responsible for delivering it to the intended recipient, you are hereby notified that you have received this message in error and that any review, dissemination, distribution or copying of this message is strictly prohibited. If you have received this communication in error, please notify us immediately by email, and delete the original message.

--
Jessica Liao | Partner Decision Scientist I | Credit Karma, Inc.
jessica.liao@creditkarma.com | www.creditkarma.com

This email may contain confidential and privileged information. Any review, use, distribution, or disclosure by anyone other than the intended recipient(s) is prohibited. If you are not the intended recipient, please contact the sender by reply email and delete all copies of this message.

--
Jessica Liao | Partner Decision Scientist I | Credit Karma, Inc.
jessica.liao@creditkarma.com | www.creditkarma.com

This email may contain confidential and privileged information. Any review, use, distribution, or disclosure by anyone other than the intended recipient(s) is prohibited. If you are not the intended recipient, please contact the sender by reply email and delete all copies of this message.

--
Luke Zhang
Senior Analyst, Risk Management
234 1st Street
San Francisco, CA 94105

SoFi

The information contained in this email message is PRIVATE and intended only for the personal and confidential use of the recipient named above. If the reader of this message is not the intended recipient or an agent responsible for delivering it to the intended recipient, you are hereby notified that you have received this message in error and that any review, dissemination, distribution or copying of this message is strictly prohibited. If you have received this communication in error, please notify us immediately by email, and delete the original message.

--
Jessica Liao | Partner Decision Scientist I | Credit Karma, Inc.
jessica.liao@creditkarma.com | www.creditkarma.com

This email may contain confidential and privileged information. Any review, use, distribution, or disclosure by anyone other than the intended recipient(s) is prohibited. If you are not the intended recipient, please contact the sender by reply email and delete all copies of this message.

--
Luke Zhang
Senior Analyst, Risk Management

234 1st Street
San Francisco, CA 94105

SoFi ⣿

The information contained in this email message is PRIVATE and intended only for the personal and confidential use of the recipient named above. If the reader of this message is not the intended recipient or an agent responsible for delivering it to the intended recipient, you are hereby notified that you have received this message in error and that any review, dissemination, distribution or copying of this message is strictly prohibited. If you have received this communication in error, please notify us immediately by email, and delete the original message.

--
Jessica Liao | Partner Decision Scientist I | Credit Karma, Inc.
jessica.liao@creditkarma.com | www.creditkarma.com

This email may contain confidential and privileged information. Any review, use, distribution, or disclosure by anyone other than the intended recipient(s) is prohibited. If you are not the intended recipient, please contact the sender by reply email and delete all copies of this message.

--
Luke Zhang
Senior Analyst, Risk Management
234 1st Street
San Francisco, CA 94105

SoFi ⣿

The information contained in this email message is PRIVATE and intended only for the personal and confidential use of the recipient named above. If the reader of this message is not the intended recipient or an agent responsible for delivering it to the intended recipient, you are hereby notified that you have received this message in error and that any review, dissemination, distribution or copying of this message is strictly prohibited. If you have received this communication in error, please notify us immediately by email, and delete the original message.

--
Jessica Liao | Partner Decision Scientist I | Credit Karma, Inc.
jessica.liao@creditkarma.com | www.creditkarma.com

This email may contain confidential and privileged information. Any review, use, distribution, or disclosure by anyone other than the intended recipient(s) is prohibited. If you are not the intended recipient, please contact the sender by reply email and delete all copies of this message.

The information contained in this email message is PRIVATE and intended only for the personal and confidential use of the recipient named above. If the reader of this message is not the intended recipient or an agent responsible for delivering it to the intended recipient, you are hereby notified that you have received this message in error and that any review, dissemination, distribution or copying of this message is strictly prohibited. If you have received this communication in error, please notify us immediately by email, and delete the original message.

--
Jessica Liao | Partner Decision Scientist I | Credit Karma, Inc.
jessica.liao@creditkarma.com | www.creditkarma.com

This email may contain confidential and privileged information. Any review, use, distribution, or disclosure by anyone other than the intended recipient(s) is prohibited. If you are not the intended recipient, please contact the sender by reply email and delete all copies of this message.

--
Jessica Liao | Partner Decision Scientist I | Credit Karma, Inc.
jessica.liao@creditkarma.com | www.creditkarma.com

--
Jessica Liao | Partner Decision Scientist I | Credit Karma, Inc.
jessica.liao@creditkarma.com | www.creditkarma.com

--
Jessica Liao | Partner Decision Scientist I | Credit Karma, Inc.
jessica.liao@creditkarma.com | www.creditkarma.com

--
Jessica Liao | Partner Decision Scientist I | Credit Karma, Inc.
jessica.liao@creditkarma.com | www.creditkarma.com

--
Luke Zhang
Senior Analyst, Risk Management
234 1st Street
San Francisco, CA 94105

SoFi

The information contained in this email message is PRIVATE and intended only for the personal and confidential use of the recipient named above. If the reader of this message is not the intended recipient or an agent responsible for delivering it to the intended recipient, you are hereby notified that you have received this message in error and that any review, dissemination, distribution or copying of this message is strictly prohibited. If you have received this communication in error, please notify us immediately by email, and delete the original message.

--
Jessica Liao | Partner Decision Scientist I | Credit Karma, Inc.
jessica.liao@creditkarma.com | www.creditkarma.com

This email may contain confidential and privileged information. Any review, use, distribution, or disclosure by anyone other than the intended recipient(s) is prohibited. If you are not the intended recipient, please contact the sender by reply email and delete all copies of this message.

--
Luke Zhang
Senior Analyst, Risk Management
234 1st Street
San Francisco, CA 94105