

DOCUMENTAÇÃO ML.BIGQUERY PANDAS FRAUD DETECTION

1. Foram realizados 3 testes:

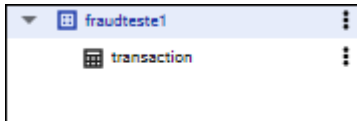
TESTE 1 com seleção de colunas diferentes.

TESTE 2 e **TESTE 2 AUTOWEIGHTS** com as mesmas colunas porém um com o **AUTO_CLASS_WEIGHTS = TRUE**.

TESTE 1

Sem o **AUTO_CLASS_WEIGHTS=TRUE**

- **IMPORTAÇÃO DA TABELA** *transaction*.



fraudteste1
transaction

Nome do campo	Tipo	Modo
ssn	STRING	NULLABLE
cc_num	INTEGER	NULLABLE
first	STRING	NULLABLE
last	STRING	NULLABLE
gender	STRING	NULLABLE
street	STRING	NULLABLE
city	STRING	NULLABLE
state	STRING	NULLABLE
zip	INTEGER	NULLABLE
city_pop	INTEGER	NULLABLE
job	STRING	NULLABLE
dob	DATE	NULLABLE
acct_num	INTEGER	NULLABLE
profile	STRING	NULLABLE
trans_num	STRING	NULLABLE
trans_date	DATE	NULLABLE
trans_time	TIME	NULLABLE
unix_time	INTEGER	NULLABLE
category	STRING	NULLABLE
amt	FLOAT	NULLABLE
is_fraud	INTEGER	NULLABLE
merchant	STRING	NULLABLE
merch_lat	STRING	NULLABLE
merch_long	STRING	NULLABLE

- **CRIANDO UMA NOVA TABELA** *is not null* DA COLUNA *is_fraud* COM AS COLUNAS SELECIONADAS.

```
CREATE TABLE `pandasfraude.fraudteste1.transaction_notnull` AS
SELECT
  ssn,
  cc_num,
  first,
  last,
  gender,
  city,
  state,
  trans_time,
  unix_time,
  amt,
  is_fraud
FROM
  `fraudteste1.transaction`
WHERE is_fraud IS NOT NULL
```

```
SELECT 'is fraud' As Category,
COUNT (*) As TotalRecords,
SUM(CASE WHEN is_fraud is null THEN 1 ELSE 0 END) AS MissingRecords,
CONCAT(CAST(ROUND(100*(SUM(CASE WHEN is_fraud is null THEN 1 ELSE 0 END)/Count(*)),2) AS STRING),'%')AS MissingRate
FROM `pandasfraude.fraudteste1.transaction_notnull`
GROUP BY 1
```

Resultados da consulta

INFORMAÇÕES DO JOB		RESULTADOS	JSON	DETALHES DA EXECUÇÃO	
Linha	Category	TotalRecords	MissingRecords	MissingRate	
1	is fraud	29301	0	0%	

COLUNAS

Nome do campo	Tipo	Modo
ssn	STRING	NULLABLE
cc_num	INTEGER	NULLABLE
first	STRING	NULLABLE
last	STRING	NULLABLE
gender	STRING	NULLABLE
city	STRING	NULLABLE
state	STRING	NULLABLE
trans_time	TIME	NULLABLE
unix_time	INTEGER	NULLABLE
amt	FLOAT	NULLABLE
is_fraud	INTEGER	NULLABLE

- TABELA **is not null**, A QUANTIDADE QUE OCORREU FRAUDE E NÃO OCORREU FRAUDE.

is_fraud = 0 → Fraude NÃO

is_fraud = 1 → Fraude SIM

```
SELECT count(*) FROM `pandasfraude.fraudteste1.transaction_notnull`
where is_fraud=0;

SELECT count(*) FROM `pandasfraude.fraudteste1.transaction_notnull`
where is_fraud=1;
```

← Resultados da consulta

INFORMAÇÕES DO JOB		RESULTADOS
Linha	f0_	
1	28668	

Resultados da consulta

INFORMAÇÕES DO JOB		RESULTADOS
Linha	f0_	
1	633	

- CRIAÇÃO DO MODELO**

A tabela possui 29301 linhas, foi retirado 15% (4395) para verificação do modelo.

```
CREATE OR REPLACE MODEL `fraudteste1.mode_log`
OPTIONS(model_type='logistic_reg') AS
SELECT
  * EXCEPT(cc_num, state, is_fraud),
  is_fraud AS label,
FROM
  `pandasfraude.fraudteste1.transaction_notnull`
ORDER BY
  first
LIMIT 24986;
```

mode_log

DETALHES

TREINAMENTO

AVALIAÇÃO

ESQUEMA

Tipo de modelo

LOGISTIC_REGRESSION

Local dos dados

us-east4

Detalhes do modelo

EDITAR

ID do modelo

pandasfraude.fraudteste1_mode_log

Descrição

Rótulos

Data de criação

Expiração do modelo

Data da modificação

Local dos dados

Tipo de modelo

LOGISTIC_REGRESSION

Tipo de perda

Perda média de registros

Dados de treinamento

TABELA TEMPORÁRIA DE DADOS DE TREINAMENTO

Dados de avaliação

TABELA TEMPORÁRIA DE DADOS DE AVALIAÇÃO

Opções de treinamento

Opções de treinamento são os parâmetros opcionais adicionados ao script para criar este modelo.

Máximo permitido de iterações

20

Iterações reais

6

Regularização de L1

0.00

Regularização de L2

0.00

mode_log

DETALHES

TREINAMENTO

AVALIAÇÃO

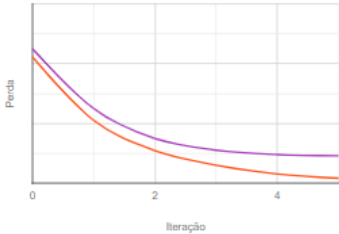
ESQUEMA

Ver como

Gráficos

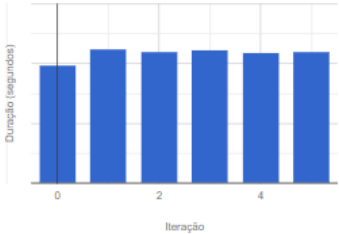
Tabela

Perda



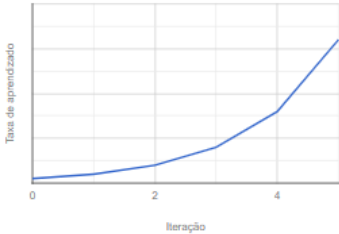
Iteração	Perda
0	0.8
1	0.6
2	0.4
3	0.3
4	0.25
5	0.2

Duração (segundos)



Iteração	Duração (segundos)
0	1.5
1	1.5
2	1.5
3	1.5
4	1.5
5	1.5

Taxa de aprendizado



Iteração	Taxa de aprendizado
0	0.001
1	0.002
2	0.005
3	0.008
4	0.012
5	0.015

mode_log

DETALHES

TREINAMENTO

AValiação

ESQUEMA

Ver como

Gráficos

Tabela

Iteração	Perda de dados de treinamento	Perda de dados de avaliação	Taxa de aprendizado	Duração (segundos)
5	0,0162	0,0916	6,4	2,20
4	0,0306	0,0956	3,2	2,19
3	0,0602	0,1103	1,6	2,23
2	0,1088	0,1498	0,8	2,20
1	0,2102	0,2492	0,4	2,24
0	0,4222	0,4500	0,2	1,97

mode_log

DETALHES

TREINAMENTO

AValiação

ESQUEMA

Métricas agregadas

Limite

0,5000

Precisão

0,4634

Retorno

0,1712

Acurácia

0,9771

Pontuação F1

0,2500

Perda de registro

0,0916

ROC AUC

0,8334

Limite de pontuação

Limite de classe positiva

0,0057

Classe positiva

1

Classe negativa

0

Precisão

0,0223

Recall

1,0000

Acurácia

0,0223

Pontuação F1

0,0436



Matriz de confusão

À tabela mostra com que frequência o modelo classificou cada rótulo corretamente (em azul) e quais foram os mais confundidos nesse rótulo (em cinza).

Rótulo verdadeiro	Rótulo previsto	
	1	0
1	100%	0%
0	100%	0%

- VIEW TRAINING AND MODEL STATISTICS

```
SELECT * FROM ML.TRAINING_INFO(MODEL `fraudteste1.mode_log`)
```

Resultados da consulta							
INFORMAÇÕES DO JOB		RESULTADOS		JSON	DETALHES DA EXECUÇÃO		
Linha	training_run	iteration	loss	eval_loss	learning_rate	duration_ms	
1	0	5	0.016165401729458927	0.091631824883518231	6.4	2203	
2	0	4	0.030611253952648219	0.095555901855801356	3.2	2188	
3	0	3	0.060152988794777	0.11028265033544002	1.6	2230	
4	0	2	0.10881465731349445	0.14977560526831324	0.8	2202	
5	0	1	0.21018916551672162	0.24918596056611081	0.4	2240	
6	0	0	0.42215021920193818	0.44995557238816242	0.2	1969	

- EVALUATION MODEL COM OS DADOS RETIRADOS (4395)

```
SELECT
*
FROM
ML.EVALUATE(
  MODEL 'fraudteste1.mode_log', (
    SELECT
      * EXCEPT(cc_num, state, is_fraud),
      is_fraud AS label,
    FROM
      'pandasfraude.fraudteste1.transaction_notnull'
    ORDER BY first DESC
    LIMIT 4395
  )
);
```

Resultados da consulta							
INFORMAÇÕES DO JOB		RESULTADOS		JSON	DETALHES DA EXECUÇÃO		
Linha	precision	recall	accuracy	f1_score	log_loss	roc_auc	
1	0.088235294117647065	0.034090909090909098	0.97360637087599544	0.049180327868852458	0.20976043393699373	0.78838061938061943	

- PREDICT

```
SELECT
*
FROM
ml.PREDICT(
  MODEL 'fraudteste1.mode_log', (
    SELECT
      *
    FROM
      'pandasfraude.fraudteste1.transaction_notnull'
  )
);
```

Resultados da consulta														
INFORMAÇÕES DO JOB		RESULTADOS			JSON	DETALHES DA EXECUÇÃO								
Linha	predicted_label	predic....label	predic....prob	ssn	cc_num	first	last	gender	city	state	trans_time	unix_time	amt	is_fraud
1	0	1	0.012520006561858758	358-97-6253	4906250000000000	Judith	Nelson	FEMININO	Bronx	NY	20:10:34	1644610234	61.15	0
		0	0.98747999343814119											
2	0	1	0.012390959629205419	358-97-6253	4906250000000000	Judith	Nelson	FEMININO	Bronx	NY	09:05:01	1650099901	63.23	0
		0	0.98760904037079456											
3	0	1	0.012815826414303295	358-97-6253	4906250000000000	Judith	Nelson	FEMININO	Bronx	NY	08:23:14	1642148594	67.8	0
		0	0.98718417358569666											
4	0	1	0.012389917166398067	358-97-6253	4906250000000000	Judith	Nelson	FEMININO	Bronx	NY	08:21:05	1645863665	59.06	0
		0	0.98761608283360192											
5	0	1	0.013049307975665664	358-97-6253	4906250000000000	Judith	Nelson	FEMININO	Bronx	NY	10:14:04	1644660844	99.86	0
		0	0.98695069202433439											
6	0	1	0.012045609691237351	358-97-6253	4906250000000000	Judith	Nelson	FEMININO	Bronx	NY	09:22:29	1644398549	8.72	0
		0	0.98795499030876264											
7	0	1	0.011725527746304918	358-97-6253	4906250000000000	Judith	Nelson	FEMININO	Bronx	NY	18:07:16	1647367636	2.65	0
		0	0.98827447225369514											
8	0	1	0.011510164349173237	358-97-6253	4906250000000000	Judith	Nelson	FEMININO	Bronx	NY	19:56:36	1650916596	13.18	0
		0	0.98848993565082671											
9	0	1	0.025186157930723461	358-97-6253	4906250000000000	Judith	Nelson	FEMININO	Bronx	NY	01:13:46	1648170826	2.17	0

- ML WEIGHTS

```
SELECT * FROM ML.WEIGHTS(MODEL `fraudteste1.model_log`)
```

	processed_input	category_weights.category	category_weights.weight
1.	trans_time	23:54:34	6.25200638291529
2.	trans_time	23:19:14	6.249403215916304
3.	trans_time	01:44:00	6.21934100281764
4.	trans_time	22:03:00	6.136214120359135
5.	trans_time	22:03:21	6.133999507024293
6.	trans_time	23:08:04	6.112043847927879
7.	trans_time	23:23:11	6.111090121846779
8.	trans_time	01:20:19	6.0983791760798765
9.	trans_time	00:08:12	6.096760996986403
10.	trans_time	01:36:19	6.0896405730321534
11.	trans_time	00:43:53	6.086323286074734
12.	trans_time	04:43:23	6.078931656462106
13.	trans_time	22:10:34	6.070104797297633
14.	trans_time	01:46:12	6.02295177447846
15.	trans_time	19:10:29	6.022913021984513
16.	trans_time	23:43:17	6.019318591913749
17.	trans_time	22:47:47	6.017813922559107
18.	trans_time	22:22:21	6.00921101653125
19.	trans_time	23:13:05	6.009170426194605
20.	trans_time	17:53:32	6.0085425439427755

TESTE 2

SEM AUTO_CLASS_WEIGHTS = TRUE

- CRIAÇÃO DA TABELA COM AS COLUNAS ESCOLHIDAS:

```
CREATE TABLE `pandasfraude.detectionfraud.fraude_notnull` AS
SELECT
  gender,
  city,
  state,
  city_pop,
  job,
  profile,
  category,
  amt,
  is_fraud,
  merchant
FROM
  `pandasfraude.detectionfraud.fraude`
WHERE is_fraud IS NOT NULL
```

fraude_notnull				
CONSULTA				
COMPARTILHAR				
ESQUEMA				
DETALHES				
VISUALIZAR				
Filtro				
Insira o nome ou o valor da propriedade				
Nome do campo	Tipo	Modo	Compilação	Tags de po
gender	INTEGER	NULLABLE		
city	STRING	NULLABLE		
state	STRING	NULLABLE		
city_pop	INTEGER	NULLABLE		
job	STRING	NULLABLE		
profile	STRING	NULLABLE		
category	STRING	NULLABLE		
amt	FLOAT	NULLABLE		
is_fraud	INTEGER	NULLABLE		
merchant	STRING	NULLABLE		

- CRIAÇÃO DO MODELO

A tabela possui 29301 linhas, foi retirado 15% (4395) para verificação do modelo.

```
CREATE OR REPLACE MODEL `detectionfraud.fraudelogreg`
OPTIONS(model_type='logistic_reg') AS
SELECT
  * EXCEPT(state, city_pop, is_fraud),
  is_fraud AS label,
FROM
  `detectionfraud.fraude_notnull`
ORDER BY
  profile
LIMIT 24906;
```

fraudelogreg

MODELO DE C

DETALHES

TREINAMENTO

AVALIAÇÃO

ESQUEMA

Tipo de modelo

LOGISTIC_REGRESSION

Local dos dados

us-east4

Detalhes do modelo

EDITAR

ID do modelo

pandasfraude:detectionfraud.fraudelogreg

Descrição

Rótulos

Data de criação

domingo, 17 de julho de 2022 17:58:00 GMT-03:00

Expiração do modelo

quinta-feira, 15 de setembro de 2022 17:58:00 GMT-03:00

Data da modificação

domingo, 17 de julho de 2022 17:58:00 GMT-03:00

Local dos dados

us-east4

Tipo de modelo

LOGISTIC_REGRESSION

Tipo de perda

Perda média de registros

Dados de treinamento

TABELA TEMPORÁRIA DE DADOS DE TREINAMENTO

Dados de avaliação

TABELA TEMPORÁRIA DE DADOS DE AVALIAÇÃO

fraudelogreg

DETALHES

TREINAMENTO

AVALIAÇÃO

ESQUEMA

Opções de treinamento

Opções de treinamento são os parâmetros opcionais adicionados ao script para criar este modelo.

Máximo permitido de iterações

20

Iterações reais

7

Regularização de L1

0,00

Regularização de L2

0,00

Parada antecipada

verdadeiro

Progresso relativo mínimo

0,01

Estratégia de taxa de aprendizado

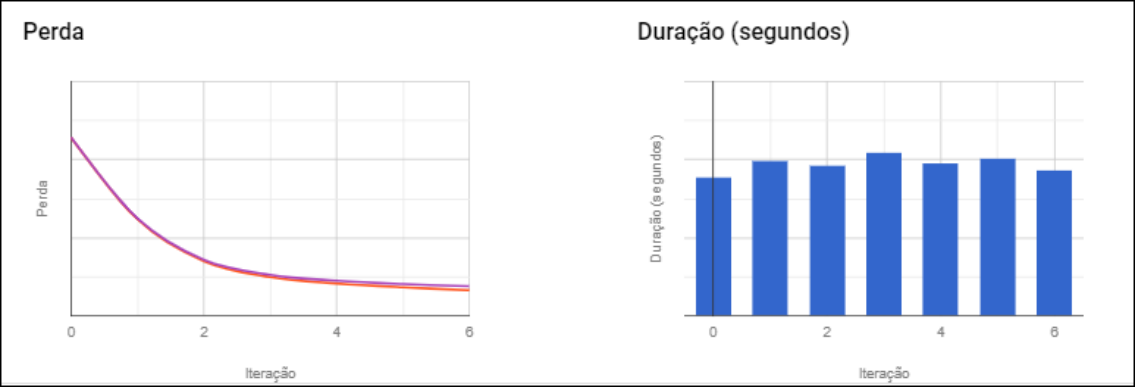
Pesquisa de linha

Taxa de aprendizagem inicial da pesquisa de linha

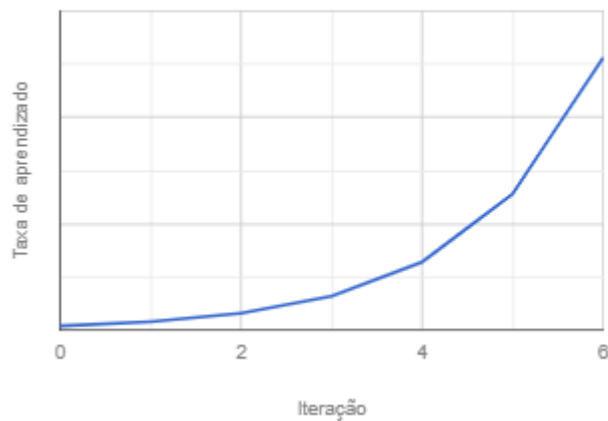
0,10

Calcular valores P

falso



Taxa de aprendizado



Ver como

- ☐ Gráficos
- ☒ Tabela

Iteração	Perda de dados de treinamento	Perda de dados de avaliação	Taxa de aprendizado	Duração (segundos)
6	0,0661	0,0762	12,8	1,86
5	0,0733	0,0818	6,4	2,01
4	0,0833	0,0904	3,2	1,95
3	0,0996	0,1051	1,6	2,10
2	0,1399	0,1438	0,8	1,92
1	0,2476	0,2499	0,4	1,99
0	0,4568	0,4577	0,2	1,77

fraudelogreg

[MODELO DE CONSULTA](#)

[EXCLUIR MODELO](#)

[EXPORT](#)

DETALHES

TREINAMENTO

AVALIAÇÃO

ESQUEMA

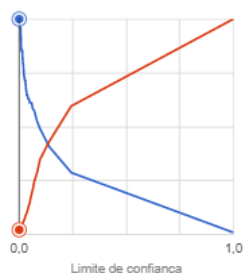
Métricas agregadas

Limite	0,5000
Precisão	0,7857
Retorno	0,1019
Acurácia	0,9801
Pontuação F1	0,1803
Perda de registro	0,0762
ROC AUC	0,8347

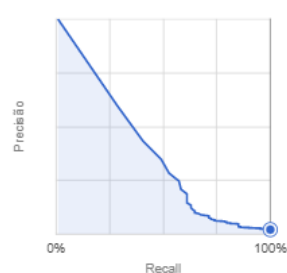
Limite de pontuação

Limite de classe positiva	<input type="text" value="0,0006"/>
Classe positiva	1
Classe negativa	0
Precisão	0,0215
Recall	1,0000
Acurácia	0,0215
Pontuação F1	0,0421

Precisão/recall por limite

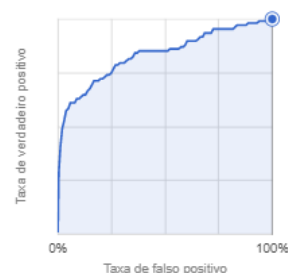


Curva de precisão/recall



Área abaixo da curva: 0,38

Curva ROC



Área abaixo da curva: 0,835

Matriz de confusão

Contagens de itens

A tabela mostra com que frequência o modelo classificou cada rótulo corretamente (em azul) e quais foram os mais confundidos nesse rótulo (em cinza).

Rótulo verdadeiro	Rótulo previsto	
	1	0
1	100%	0%
0	100%	0%

- VIEW TRAINING AND MODEL STATISTICS

```
SELECT * FROM ML.TRAINING_INFO(MODEL `detectionfraud.fraudelogreg`)
```

Resultados da consulta

SALVAR

INFORMAÇÕES DO JOB		RESULTADOS		JSON	DETALHES DA EXECUÇÃO	
Linha	training_run	iteration	loss	eval_loss	learning_rate	duration_ms
1	0	6	0.0660689048175093	0.076211590107441288	12.8	1862
2	0	5	0.0733322592129123	0.0818471968291359	6.4	2013
3	0	4	0.083274893248983586	0.090380551469115028	3.2	1953
4	0	3	0.099596756583153048	0.10514496604271718	1.6	2102
5	0	2	0.13992833832035029	0.14380429430332167	0.8	1924
6	0	1	0.24756823850895485	0.2498506750300602	0.4	1994
7	0	0	0.45678170962249204	0.45771485850754284	0.2	1772

- EVALUATION MODEL

```
SELECT
*
FROM
ML.EVALUATE(
  MODEL `detectionfraud.fraudelogreg`, (
    SELECT
      * EXCEPT(state, city_pop, is_fraud),
      is_fraud AS label,
    FROM
      `detectionfraud.fraude_notnull`
    ORDER BY profile DESC
    LIMIT 4395
  )
);
```

Resultados da consulta

SALVAR RESULTADOS EXPLORAR D

INFORMAÇÕES DO JOB		RESULTADOS		JSON	DETALHES DA EXECUÇÃO	
Linha	precision	recall	accuracy	f1_score	log_loss	roc_auc
1	0.5168539325842697	0.32857142857142857	0.96882821387940843	0.40174672489082969	0.270673729250454	0.86228471528471529

- PREDICT

```
SELECT
*
FROM
ml.PREDICT(
  MODEL `detectionfraud.fraudelogreg`, (
    SELECT
      * EXCEPT(state, city_pop)
    FROM
      `detectionfraud.fraude_notnull`
  )
);
```

Resultados da consulta

SALVAR RESULTADOS

EXPLO

INFORMAÇÕES DO JOB			RESULTADOS	JSON	DETALHES DA EXECUÇÃO							
Linha	predicted_label	predic...label	predic...prob	gender	city	job	profile	category	amt	is_fraud	merchant	
1	0	1	0.024193180236516402	FEMININO	Bronx	Chartered accountant	adults_50up_female_urban.json	home	159.47	0	fraud_Schmidt-Larkin	
		0	0.9758068197634836									
2	0	1	0.016103652435063971	FEMININO	Bronx	Chartered accountant	adults_50up_female_urban.json	home	9.37	0	fraud_Durgan	
		0	0.983896347564936									
3	0	1	0.016563050228628664	FEMININO	Bronx	Chartered accountant	adults_50up_female_urban.json	home	17.55	0	fraud_Stark-Koss	
		0	0.9834369497713713									
4	0	1	0.0167859938660795	FEMININO	Bronx	Chartered accountant	adults_50up_female_urban.json	home	25.86	0	fraud_Emsen-Feest	
		0	0.98321400613392052									

• EXPLAIN MODEL

```
SELECT * FROM ML.WEIGHTS(MODEL `detectionfraud.fraudelogreg`)
```

	processed_input	category_weights.category	category_weights.weight
1.	city	East Liverpool	6.020137176654214
2.	job	Music tutor	3.5396194484601717
3.	city	Kingsport	3.5396194484601717
4.	job	Private music teacher	3.1673547046539454
5.	city	Austin	3.1673547046539454
6.	job	Community education officer	2.968887100682255
7.	city	Orlando	2.968887100682255
8.	merchant	fraud_Gibson-Deckow	1.7844888200591091
9.	merchant	fraud_Hamill-Daugherty	1.7447740745286155
10.	merchant	fraud_Stehr	1.5338814946022903
11.	merchant	fraud_Huel-Langworth	1.4614119752266637
12.	merchant	fraud_Padberg-Welch	1.3109856792099994
13.	merchant	fraud_Rodriguez	1.2955908989732294
14.	merchant	fraud_Dooley-Thompson	1.2225352378952936
15.	merchant	fraud_Kuhic LLC	1.1398444020336576
16.	merchant	fraud_Hayes	0.9950717838359999
17.	merchant	fraud_Fritsch LLC	0.9512126416586079
18.	merchant	fraud_Heathote LLC	0.848004186166883
19.	merchant	fraud_Kassulke PLC	0.8456628318786367
20.	merchant	fraud_Berge LLC	0.8159448328275194

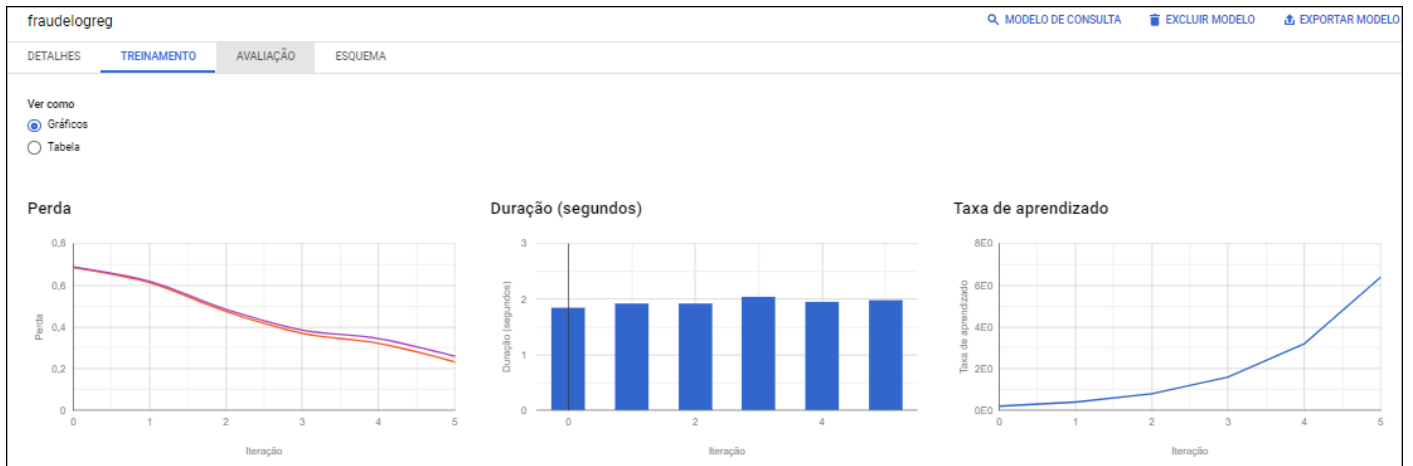
TESTE 2

COM AUTO_CLASS_WEIGHTS = TRUE

- CRIAÇÃO DO MODELO

A tabela possui 29301 linhas, foi retirado 15% (4395) para verificação do modelo.

```
CREATE OR REPLACE MODEL 'detectionfraud.fraudelogreg'  
OPTIONS(model_type='logistic_reg', AUTO_CLASS_WEIGHTS=TRUE) AS  
SELECT  
  * EXCEPT(state, city_pop, is_fraud),  
  is_fraud AS label,  
FROM  
  'detectionfraud.fraude_notnull'  
ORDER BY  
  profile  
LIMIT 24986;
```



fraudelogreg

DETALHES TREINAMENTO AVALIAÇÃO ESQUEMA

Ver como

☐ Gráficos ☒ Tabela

Iteração	Perda de dados de treinamento	Perda de dados de avaliação	Taxa de aprendizado	Duração (segundos)
5	0,2327	0,2590	6,4	1,99
4	0,3216	0,3445	3,2	1,96
3	0,3695	0,3849	1,6	2,04
2	0,4749	0,4851	0,8	1,93
1	0,6124	0,6188	0,4	1,93
0	0,6848	0,6877	0,2	1,85

fraudelogreg

MODEL DE CONSULTA EXCLUIR MODELO

DETALHES TREINAMENTO AVALIAÇÃO ESQUEMA

Métricas agregadas

Limite	0,5000
Precisão	0,1119
Retorno	0,6566
Acurácia	0,8894
Pontuação F1	0,1912
Perda de registro	0,2590
ROC AUC	0,8496

Limite de pontuação

Limite de classe positiva 0,0024

Classe positiva 1

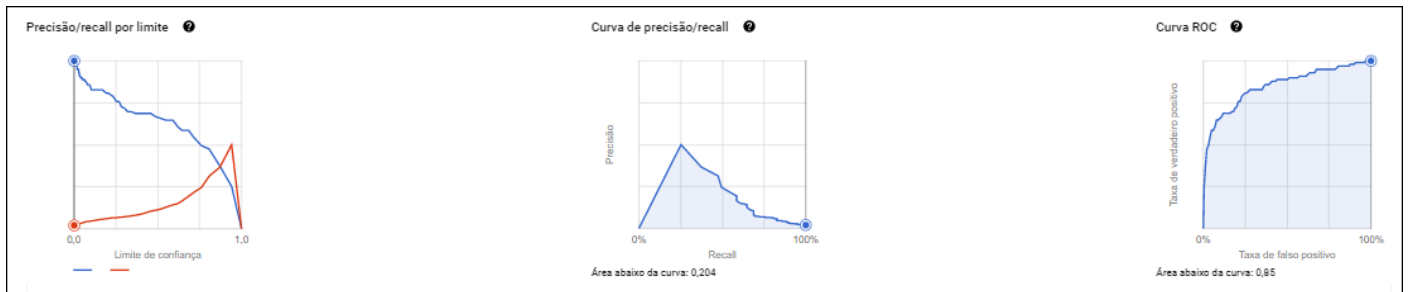
Classe negativa 0

Precisão 0,0199

Recall 1,0000

Acurácia 0,0199

Pontuação F1 0,0390



Matriz de confusão

A tabela mostra com que frequência o modelo classificou cada rótulo corretamente (em azul) e quais foram os mais confundidos nesse rótulo (em cinza).

Rótulo verdadeiro	Rótulo previsto	
	1	0
1	100%	0%
0	100%	0%

VIEW TRAINING AND MODEL STATISTICS

```
SELECT * FROM ML.TRAINING_INFO(MODEL `detectionfraud.fraudelogreg`)
```

Resultados da consulta

INFORMAÇÕES DO JOB		RESULTADOS		JSON	DETALHES DA EXECUÇÃO	
Linha	training_run	iteration	loss	eval_loss	learning_rate	duration_ms
1	0	5	0.23273609039595519	0.25900120443167751	6.4	1994
2	0	4	0.32155541574638391	0.34454977382012533	3.2	1964
3	0	3	0.36948390905071443	0.38494275776209708	1.6	2044
4	0	2	0.47493651140184023	0.48512749872368538	0.8	1927
5	0	1	0.61243352305328	0.618797773392912	0.4	1925
6	0	0	0.6848020869916428	0.68766665068281274	0.2	1847

EVALUATION MODEL

```
SELECT
| *
FROM
  ML.EVALUATE(
    MODEL `detectionfraud.fraudelogreg`, (
      SELECT
        * EXCEPT(state, city_pop, is_fraud),
        is_fraud AS label,
      FROM
        `detectionfraud.fraude_notnull`
      ORDER BY profile DESC
      LIMIT 4395
    )
  );
```

Resultados da consulta

INFORMAÇÕES DO JOB		RESULTADOS		JSON	DETALHES DA EXECUÇÃO	
Linha	precision	recall	accuracy	f1_score	log_loss	roc_auc
1	0.14682080924855492	0.92028985507246375	0.92957906712172924	0.25324027916251246	0.5177075076623544	0.9331038961038961

PREDICT

```

SELECT
  *
FROM
  ml.PREDICT(
    MODEL 'detectionfraud.fraudeologreg', (
      SELECT
        * EXCEPT(state, city_pop)
      FROM
        'detectionfraud.fraude_notnull'
    )
  );

```

Resultados da consulta													
INFORMAÇÕES DO JOB			RESULTADOS	JSON	DETALHES DA EXECUÇÃO								
Linha	predicted_label	predic...label	predic...prob	gender	city	job	profile	category	amt	is_fraud	merchant	SPLIT	
1	0	1	0.26381331605130726	FEMININO	Bronx	Chartered accountant	adults_50up_female_urban.json	home	159.47	0	fraud_Schmidt-Larkin	null	
		0	0.7361866839486928										
2	0	1	0.16142930005134379	FEMININO	Bronx	Chartered accountant	adults_50up_female_urban.json	home	9.37	0	fraud_Dungan	null	
		0	0.83857069994865618										
3	0	1	0.166192607572752	FEMININO	Bronx	Chartered accountant	adults_50up_female_urban.json	home	17.55	0	fraud_Stark-Koss	null	
		0	0.833807392427248										
4	0	1	0.16966022342038306	FEMININO	Bronx	Chartered accountant	adults_50up_female_urban.json	home	25.86	0	fraud_Ernsen-Feest	null	
		0	0.830339776579617										
5	0	1	0.17837736778130703	FEMININO	Bronx	Chartered accountant	adults_50up_female_urban.json	home	39.96	0	fraud_Ernsen-Feest	null	
		0	0.821622632218693										

- EXPLAIN MODEL

```

SELECT * FROM ML.WEIGHTS(MODEL 'detectionfraud.fraudeologreg')

```

	processed_input	category_weights.category	category_weights.weight
1.	city	East Liverpool	5.799209691878422
2.	job	Music tutor	3.7903394656780254
3.	city	Kingsport	3.7903394656780254
4.	merchant	fraud_Reichert-Weissnat	3.6682076568208766
5.	merchant	fraud_Kozey-Kuhlman	3.5257847377600204
6.	job	Private music teacher	3.405738282956069
7.	city	Austin	3.405738282956069
8.	merchant	fraud>Weimann-Lockman	3.3893589250179588
9.	job	Community education officer	3.3324095525863258
10.	city	Orlando	3.3324095525863258
11.	merchant	fraud_Reilly and Sons	3.2359065956509916
12.	merchant	fraud_Schiller Ltd	3.225530044655185
13.	merchant	fraud_Block Group	3.1121491026460464
14.	merchant	fraud_Labadie LLC	3.0453683295998317
15.	merchant	fraud_Connelly PLC	2.893895070630142
16.	merchant	fraud_Ritchie	2.8659997232857064
17.	merchant	fraud_Lakin	2.7888853935472215
18.	merchant	fraud_Gottlieb-Hansen	2.7883877146115874
19.	merchant	fraud_Nader-Maggio	2.704018678695351
20.	merchant	fraud_Funk Group	2.6795635315665534