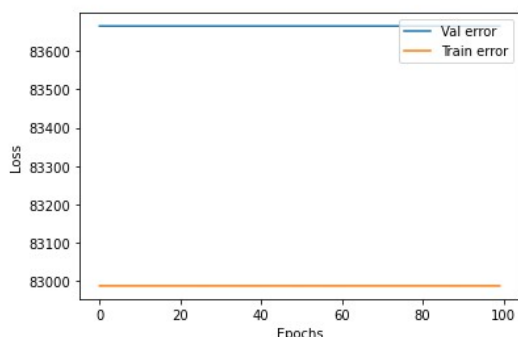


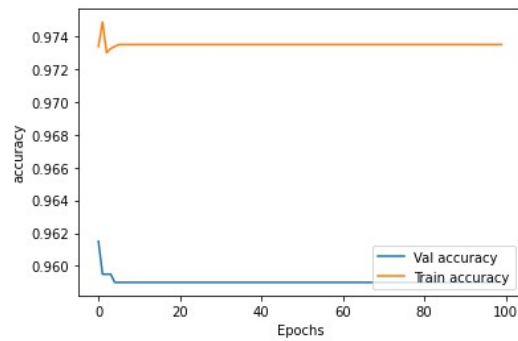
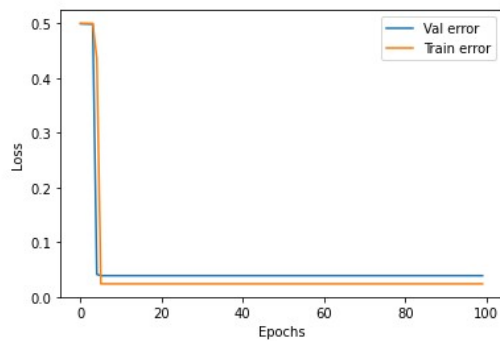
راه حل و ایده کلی

هدف از این پروژه به کار گیری اتوانکدر جهت تعیین جعلی بودن یا نبودن تراکنش های مالی انجام شده با توجه دیتا داده شده مربوط به هر فرد است. برای این کار ما یکبار با استفاده از دیتای آیدنتیتی به همراه دیتای ترنزکشن و یکبار فقط دیتای ترنزکشن آموزش مدل را انجام داده ایم. روش کلی کار به این صورت است که ابتدا داده های نال موجود در دیتاست داده شده را پر میکنیم و لیبل انکودینگ انجام می دهیم و سپس دیتا را نرمال سازی کرده و مرحله بعدی که طراحی مدل میباشد را شروع میکنیم. طراحی مدل این صورت است که ما اتوانکدر را با استفاده از داده های آموزشی ترین میکنیم تا به بهترین وزن ها برسیم سپس لیتنت وکتور را با استفاده از انکودر بدست می آوریم در محله بعد لیتنت وکتور را به همراه لیبل داده های آموزشی به کلسیفایر می دهیم تا لرن بشه و در نهایت از بهترین مدل بدست آمده برای پیش بینی خروجی داده های تست استفاده میکنیم

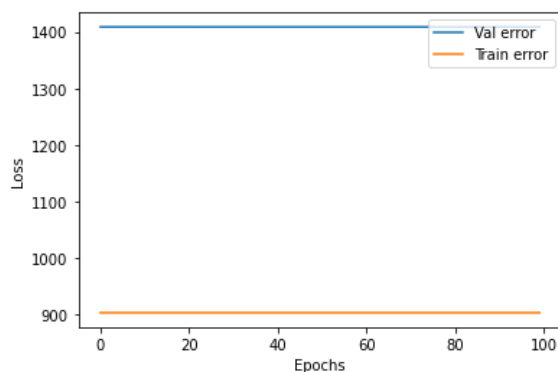
ارزیابی نتایج



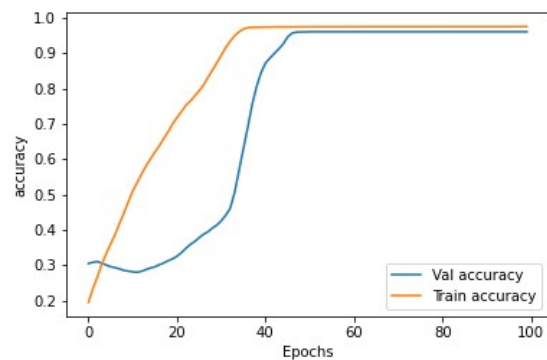
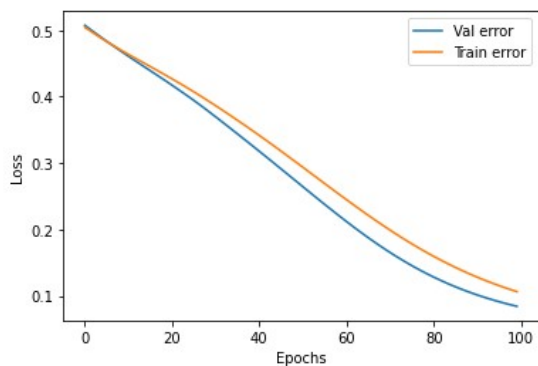
این نمودار لاس مربوط به اتوانکدر است که با استفاده از داده های آیدنتیتی و ترنزکشن طراحی شده است در اینجا از آپتیمایزر آدام استفاده شده و تعداد اپاک ها ۱۰۰ و بتچ سایز ۳۰۰ و لرنینگ ریت برابر (۰.۲۰) ۱۰۸ است آپتیمایزر را به اس جی دی هم تغییر دادیم و تعداد داده هایی را که برای لرن شدن به سیستم می دهیم و تعداد اپاک ها و بتچ سایز و لرنینگ ریت را نیز تغییر دادیم ولی نتیجه ای حاصل نشد (زمان خوبی هم برای این قسمت صرف شد)



این دو نمودار مربوط به مدل کلسیفایر (با استفاده از داده های آیدنتیتی و ترنزکشن) می باشند طراحی این مدل به این صورت است که از دو لایه هیدن تشکیل شده و تابع فعالسازی استفاده شده رلو است در ادامه از آپتیمایزر اس جی دی با لرنینگ ریت 0.0001 استفاده کردیم و بتج سائز و تعداد ایپاک ها به ترتیب ۱۲۸ و ۱۰۰ می باشند و از مین ابلوت ارور به عنوان تابع لاس استفاده شده است دقت بدست آمده از اجرای مدل روی داده های ولیدیشن (قسمتی از داده های ترین را به عنوان ولیدیشن جدا کردیم)، حدوداً ۹۶ درصد است



این نمودار لاس مربوط به اتوانکدر است که با استفاده از داده های ترنزکشن طراحی شده است تنها تفاوتش با قبلی در این است که از دیتای آیدنتیتی استفاده نکردیم



این دو نمودار مربوط به مدل کلسیفایر (با استفاده از داده ها ترنزشن) می باشند و باقی مدل مانند حالت قبل طراحی شده است

Model: "Encoder"

Layer (type)	Output Shape	Param #
input_encoder (InputLayer)	[(None, 226)]	0
encoder_layer1 (Dense)	(None, 1024)	232448
encoder_layer2 (Dense)	(None, 512)	524800
encoder_layer3 (BatchNormali	(None, 512)	2048
encoder_layer4 (Dropout)	(None, 512)	0
encoder_layer5 (Dense)	(None, 256)	131328
encoder_layer6 (Dense)	(None, 128)	32896
encoder_layer7 (Dense)	(None, 128)	16512
encoder_layer8 (Dense)	(None, 64)	8256
encoder_layer9 (Dense)	(None, 64)	4160
output_encoder (Dense)	(None, 64)	4160
Total params: 956,608		
Trainable params: 955,584		
Non-trainable params: 1,024		

Model: "Decoder"

Layer (type)	Output Shape	Param #
input_decoder (InputLayer)	[(None, 64)]	0
decoder_layer1 (Dense)	(None, 64)	4160
decoder_layer2 (Dense)	(None, 64)	4160
decoder_layer3 (BatchNormali	(None, 64)	256
decoder_layer4 (Dropout)	(None, 64)	0
decoder_layer5 (Dense)	(None, 128)	8320
decoder_layer6 (Dense)	(None, 128)	16512
decoder_layer7 (Dense)	(None, 256)	33024
decoder_layer8 (Dense)	(None, 512)	131584
decoder_layer9 (Dense)	(None, 1024)	525312
output_layer (Dense)	(None, 226)	231650
Total params: 954,978		
Trainable params: 954,850		
Non-trainable params: 128		

مدل انکودر و دیکودر که با استفاده از دیتای آیدنتیتی و ترنزشن طراحی کردیم که هرکدام از ۷ لایه هیدن تشکیل شده اند و تابع فعالسازی به کار رفته رلو می باشد

```
Model: "AutoEncoder"
```

```
-----  
Layer (type)                Output Shape              Param #  
-----  
input_autoencoder (InputLay [None, 226])              0  
-----  
Encoder (Functional)         (None, 64)                956608  
-----  
Decoder (Functional)         (None, 226)               954978  
-----  
Total params: 1,911,586  
Trainable params: 1,910,434  
Non-trainable params: 1,152  
-----
```

اتوانکدر طراحی شده با استفاده از دیتای آیدنتیتی و ترنزکشن.

```
Epoch 97/100  
27/27 [=====] - 1s 55ms/step - loss: 82988.1250 - val_loss: 83664.  
3281  
Epoch 98/100  
27/27 [=====] - 1s 49ms/step - loss: 82988.1250 - val_loss: 83664.  
3281  
Epoch 99/100  
27/27 [=====] - 1s 51ms/step - loss: 82988.1172 - val_loss: 83664.  
3281  
Epoch 100/100  
27/27 [=====] - 1s 49ms/step - loss: 82988.1328 - val_loss: 83664.  
3281
```

لاس ۸۲۹۸۸ است (عدد نجومی!)
همانطور هم که بالاتر اشاره شد تغییراتی را اعمال کردیم ولی نتیجه ای حاصل نشد (با داده های
ترنزکشن و آیدنتیتی)

```
Model: "functional_1"
```

Layer (type)	Output Shape	Param #
input_layer (InputLayer)	[(None, 64)]	0
dense (Dense)	(None, 64)	4160
dense_1 (Dense)	(None, 32)	2080
dense_2 (Dense)	(None, 2)	66

Total params: 6,306
Trainable params: 6,306
Non-trainable params: 0

مدل کلسیفایر طراحی شده با استفاده از دیتای آیدنتیتی و ترنزکشن که از دو لایه هیدن تشکیل شده است

```
Epoch 93/100
63/63 [=====] - 0s 2ms/step - loss: 0.0237 - accuracy: 0.9735 - val_loss: 0.0386 - val_accuracy: 0.9598
Epoch 94/100
63/63 [=====] - 0s 2ms/step - loss: 0.0237 - accuracy: 0.9735 - val_loss: 0.0386 - val_accuracy: 0.9598
Epoch 95/100
63/63 [=====] - 0s 2ms/step - loss: 0.0237 - accuracy: 0.9735 - val_loss: 0.0386 - val_accuracy: 0.9598
Epoch 96/100
63/63 [=====] - 0s 2ms/step - loss: 0.0237 - accuracy: 0.9735 - val_loss: 0.0386 - val_accuracy: 0.9598
Epoch 97/100
63/63 [=====] - 0s 2ms/step - loss: 0.0237 - accuracy: 0.9735 - val_loss: 0.0386 - val_accuracy: 0.9598
Epoch 98/100
63/63 [=====] - 0s 2ms/step - loss: 0.0237 - accuracy: 0.9735 - val_loss: 0.0386 - val_accuracy: 0.9598
Epoch 99/100
63/63 [=====] - 0s 2ms/step - loss: 0.0237 - accuracy: 0.9735 - val_loss: 0.0386 - val_accuracy: 0.9598
Epoch 100/100
63/63 [=====] - 0s 2ms/step - loss: 0.0237 - accuracy: 0.9735 - val_loss: 0.0386 - val_accuracy: 0.9598
```

بعد از اجرای مدل با ۱۰۰ اپیاک به دقت ۹۷٪ میرسیم
(دیتای آیدنتیتی و ترنزکشن استفاده شده است)

```

Epoch 97/100
27/27 [=====] - 1s 38ms/step - loss: 903.9241 - val_loss: 1408.545
4
Epoch 98/100
27/27 [=====] - 1s 38ms/step - loss: 903.9241 - val_loss: 1408.545
4
Epoch 99/100
27/27 [=====] - 1s 40ms/step - loss: 903.9240 - val_loss: 1408.545
4
Epoch 100/100
27/27 [=====] - 1s 38ms/step - loss: 903.9242 - val_loss: 1408.545
4

```

لاس با مقدار 903 که برای مدلی است که با داده های ترنزکشن طراحی شده است

```

Epoch 93/100
63/63 [=====] - 0s 2ms/step - loss: 0.0934 - accuracy: 0.9760 - val_loss: 0.0652 - val_accuracy: 0.9615
Epoch 94/100
63/63 [=====] - 0s 2ms/step - loss: 0.0916 - accuracy: 0.9760 - val_loss: 0.0642 - val_accuracy: 0.9615
Epoch 95/100
63/63 [=====] - 0s 2ms/step - loss: 0.0899 - accuracy: 0.9760 - val_loss: 0.0632 - val_accuracy: 0.9615
Epoch 96/100
63/63 [=====] - 0s 2ms/step - loss: 0.0883 - accuracy: 0.9760 - val_loss: 0.0623 - val_accuracy: 0.9615
Epoch 97/100
63/63 [=====] - 0s 2ms/step - loss: 0.0867 - accuracy: 0.9760 - val_loss: 0.0614 - val_accuracy: 0.9615
Epoch 98/100
63/63 [=====] - 0s 2ms/step - loss: 0.0852 - accuracy: 0.9760 - val_loss: 0.0606 - val_accuracy: 0.9615
Epoch 99/100
63/63 [=====] - 0s 2ms/step - loss: 0.0837 - accuracy: 0.9760 - val_loss: 0.0597 - val_accuracy: 0.9615
Epoch 100/100
63/63 [=====] - 0s 2ms/step - loss: 0.0823 - accuracy: 0.9760 - val_loss: 0.0590 - val_accuracy: 0.9615

```

بعد از اجرای مدل با ۱۰۰ اپیاک به دقت ۹۷٪ میرسیم
(دیتای ترنزکشن استفاده شده است)

مسئله ای که ما اینجا داشتیم نوعی آنومالی دیتکشن است. و یجورایی کلسیفیکیشن محسوب میشود
برای حل این مسئله ما از اتوانکدر استفاده کردیم و ساختاری که در این تمرین به کار بردیم یک
ساختار سمی سوپروایزد است (حالتی بین سوپروایزد و آنسوپروایزد)