

T-Distributed Stochastic Neighbor Embedding

1. t-SNE on boundary values

- ##### CASE-I
- ##### CASE-II
- ##### CASE-III

2. Multiple runs of t-SNE on range of perplexities and iterations

- ##### CASE-I
- ##### CASE-II
- ##### CASE-III

3. t-SNE on Iris Dataset

- ##### CASE-I
- ##### CASE-II
- ##### CASE-III
- ##### Results from Multiple Runs
- ##### t-SNE with 'pca' embedding

4. Conclusion

In [1]:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

from sklearn.datasets import load_breast_cancer as bc, load_iris as l_iris
from sklearn.preprocessing import StandardScaler as SS
from sklearn.manifold import TSNE

%matplotlib inline
```

In [2]:

```
cancer = bc()                      ## Instantiating Breast Cancer Database
iris = l_iris()                     ## Instantiating IRIS Dataset object
ss = SS()                           ## Instantiating Standard Scaler
tsne = TSNE(n_components=2, random_state=41) ## Instantiating TSNE
```

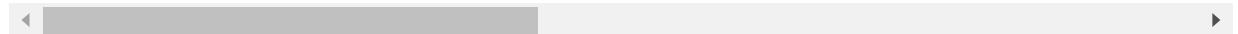
In [3]:

```
cancer_df = pd.concat([pd.DataFrame(cancer.data, columns=cancer.feature_names), pd.Dat
cancer_norm_df = pd.DataFrame(ss.fit_transform(cancer_df.iloc[:,0:-1]), columns=cance
cancer_norm_df.head(10)
```

Out[3]:

	mean radius	mean texture	mean perimeter	mean area	mean smoothness	mean compactness	mean concavity	mean concave points	n symm
0	1.097064	-2.073335	1.269934	0.984375	1.568466	3.283515	2.652874	2.532475	2.21
1	1.829821	-0.353632	1.685955	1.908708	-0.826962	-0.487072	-0.023846	0.548144	0.00
2	1.579888	0.456187	1.566503	1.558884	0.942210	1.052926	1.363478	2.037231	0.93
3	-0.768909	0.253732	-0.592687	-0.764464	3.283553	3.402909	1.915897	1.451707	2.86
4	1.750297	-1.151816	1.776573	1.826229	0.280372	0.539340	1.371011	1.428493	-0.00
5	-0.476375	-0.835335	-0.387148	-0.505650	2.237421	1.244335	0.866302	0.824656	1.00
6	1.170908	0.160649	1.138125	1.095295	-0.123136	0.088295	0.300072	0.646935	-0.06
7	-0.118517	0.358450	-0.072867	-0.218965	1.604049	1.140102	0.061026	0.281950	1.40
8	-0.320167	0.588830	-0.184080	-0.384207	2.201839	1.684010	1.219096	1.150692	1.96
9	-0.473535	1.105439	-0.329482	-0.509063	1.582699	2.563358	1.738872	0.941760	0.79

10 rows × 30 columns



In [4]: cancer_df['target'].value_counts()

Out[4]: 1 357
0 212
Name: target, dtype: int64

t-SNE_on_boundary_values

Boundary: CASE-I

- ### Lowest boundary value of Perplexity as 1 on different Iterations

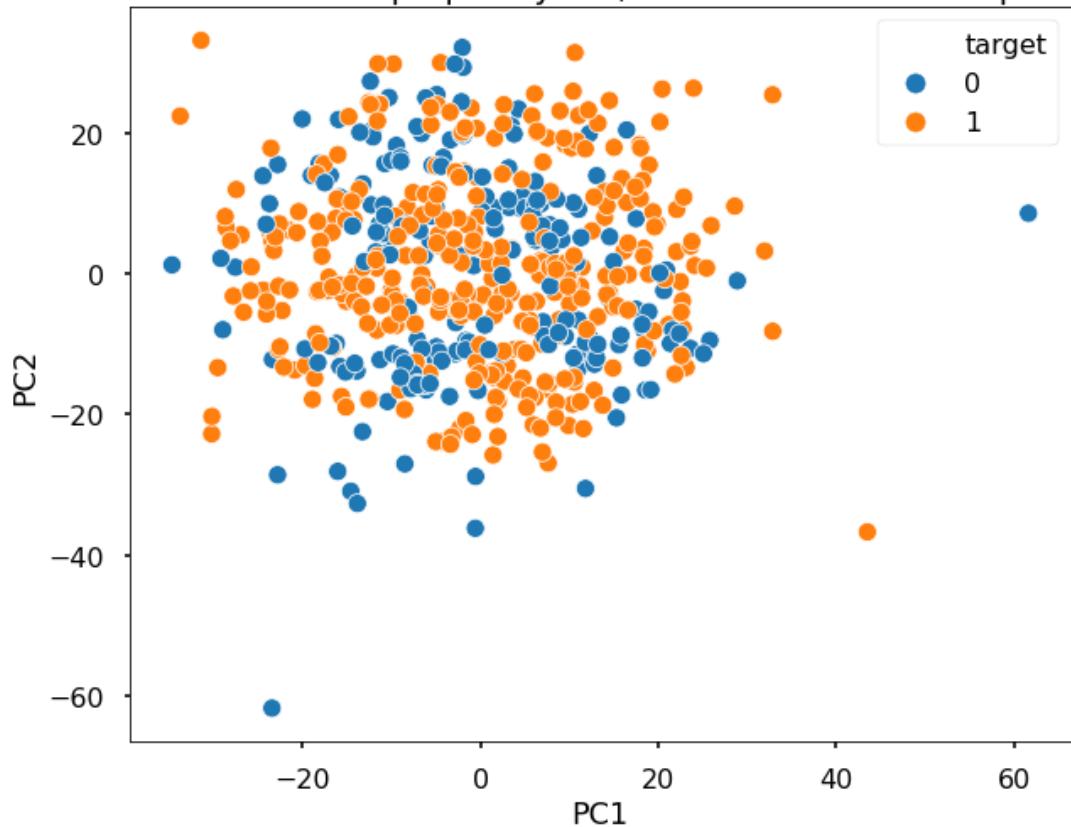
In [83]: iterations = [250, 350, 500, 750, 1000, 2500, 3500, 4000, 5000]

```

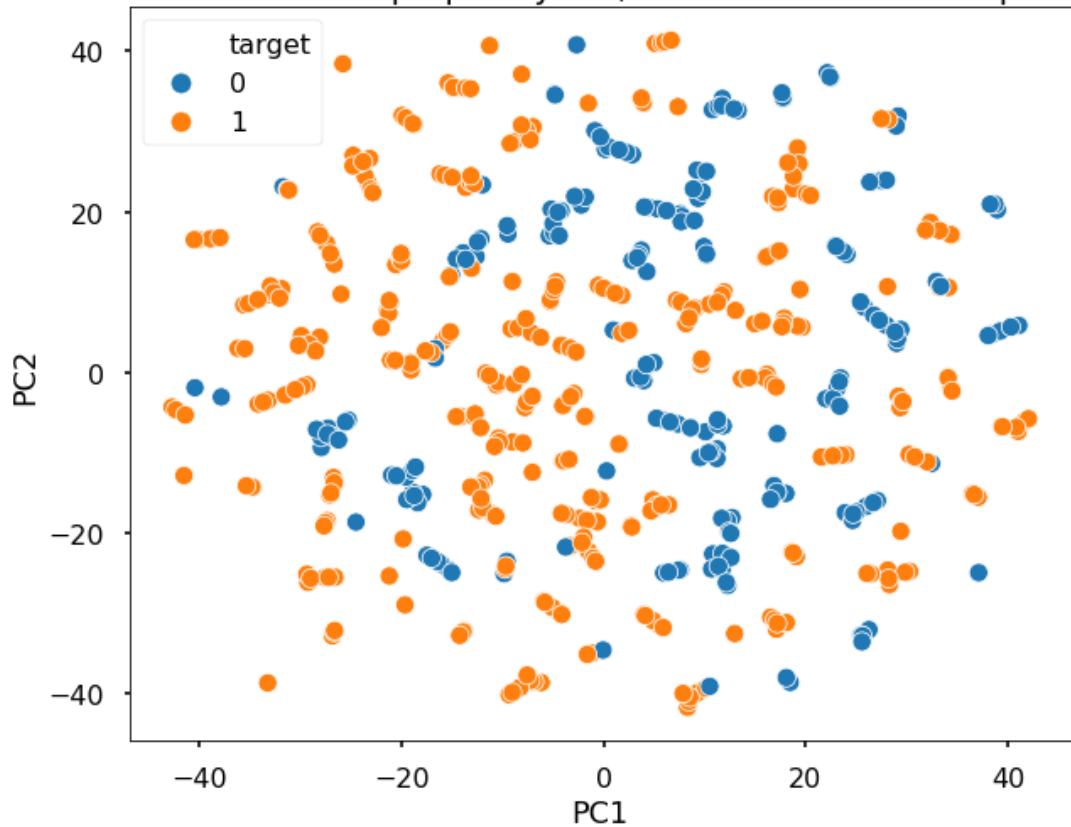
for idx in range(len(iterations)):
    tsne1 = TSNE(n_components=2, perplexity=1, learning_rate=200, n_iter=iterations[idx])
    cancer_tsne_pcmps = pd.DataFrame(tsne1.fit_transform(cancer_norm_df), columns=['PC1', 'PC2'])
    cancer_tsne_pcmps = pd.concat([cancer_tsne_pcmps, cancer_df['target']], axis=1)
    with plt.style.context('seaborn-poster'):
        plt.figure(figsize=(7,7))
        sns.scatterplot(data=cancer_tsne_pcmps, x='PC1', y='PC2', hue='target')
        plt.title("t-SNE visualization with perplexity -- {0}, iterations -- {1} and".format(1, iterations[idx]))
    plt.show()

```

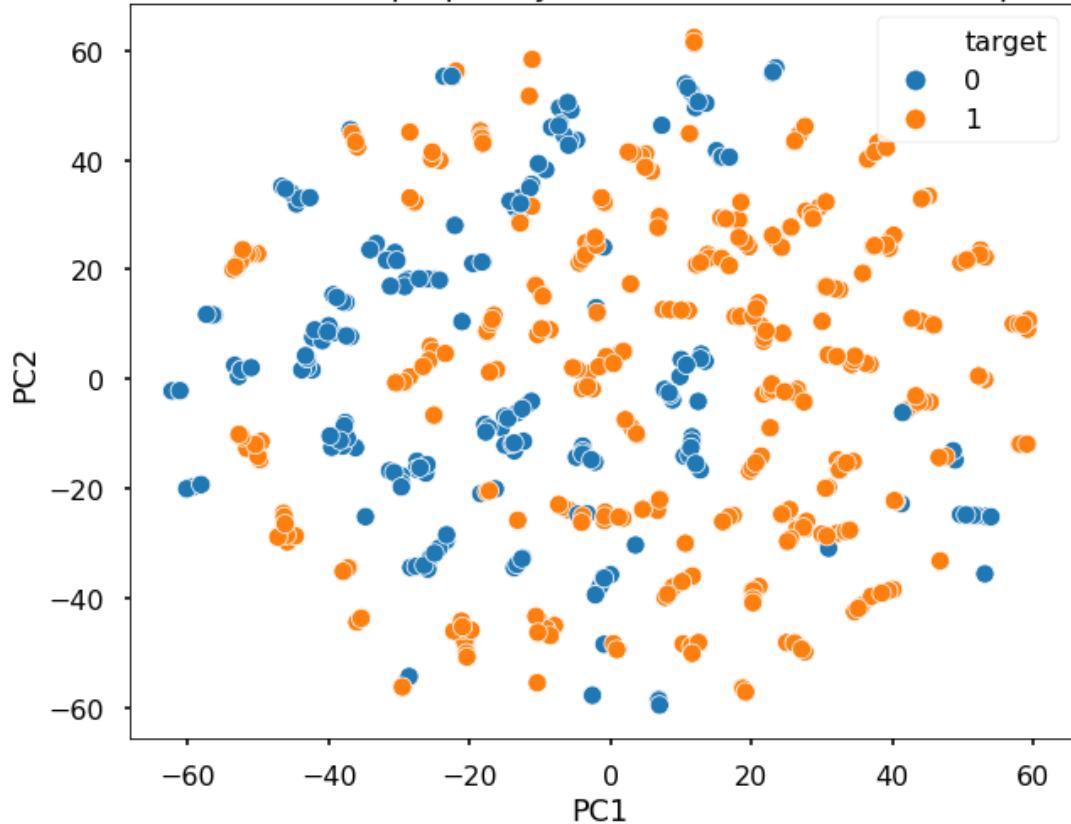
t-SNE visualization with perplexity -- 1, iterations -- 250 and epsilon -- 200



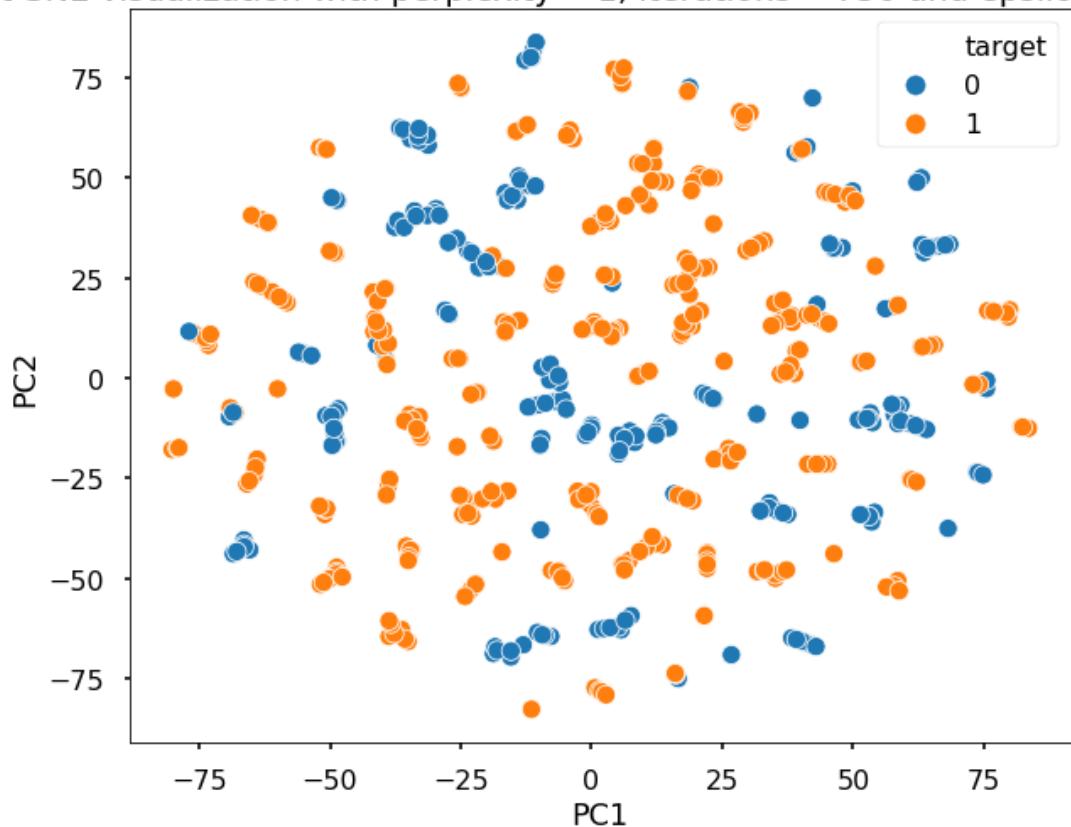
t-SNE visualization with perplexity -- 1, iterations -- 350 and epsilon -- 200



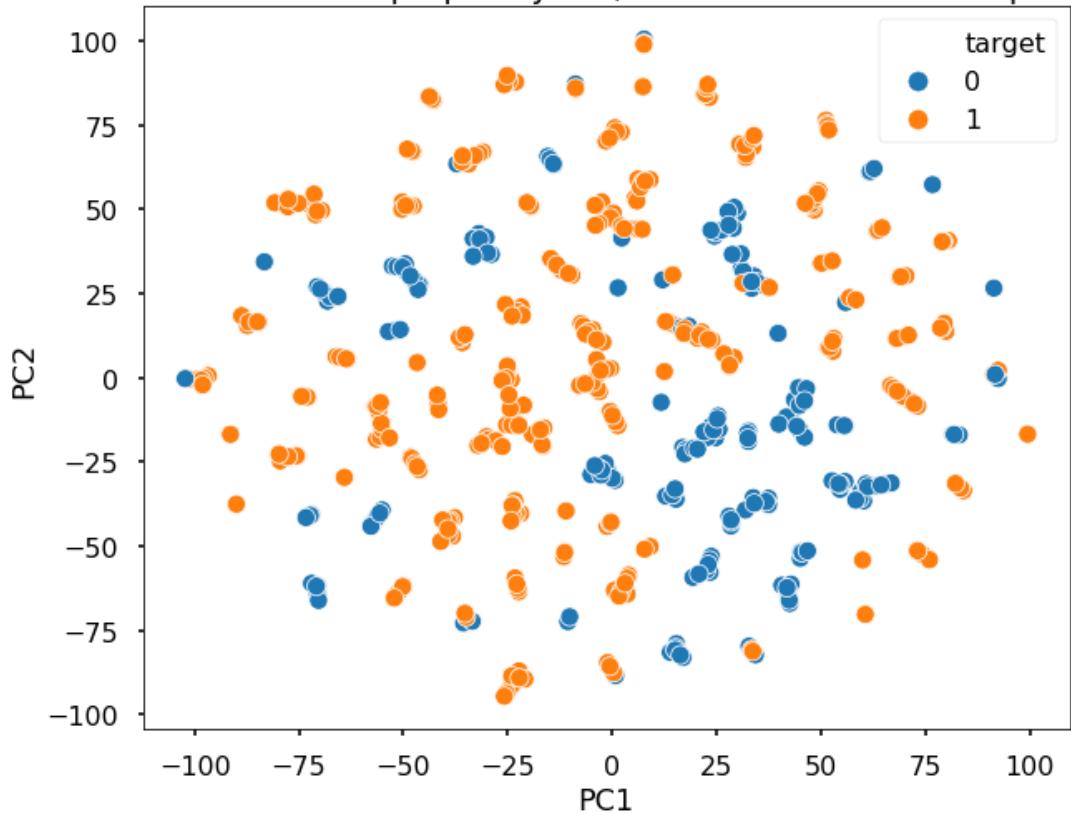
t-SNE visualization with perplexity -- 1, iterations -- 500 and epsilon -- 200



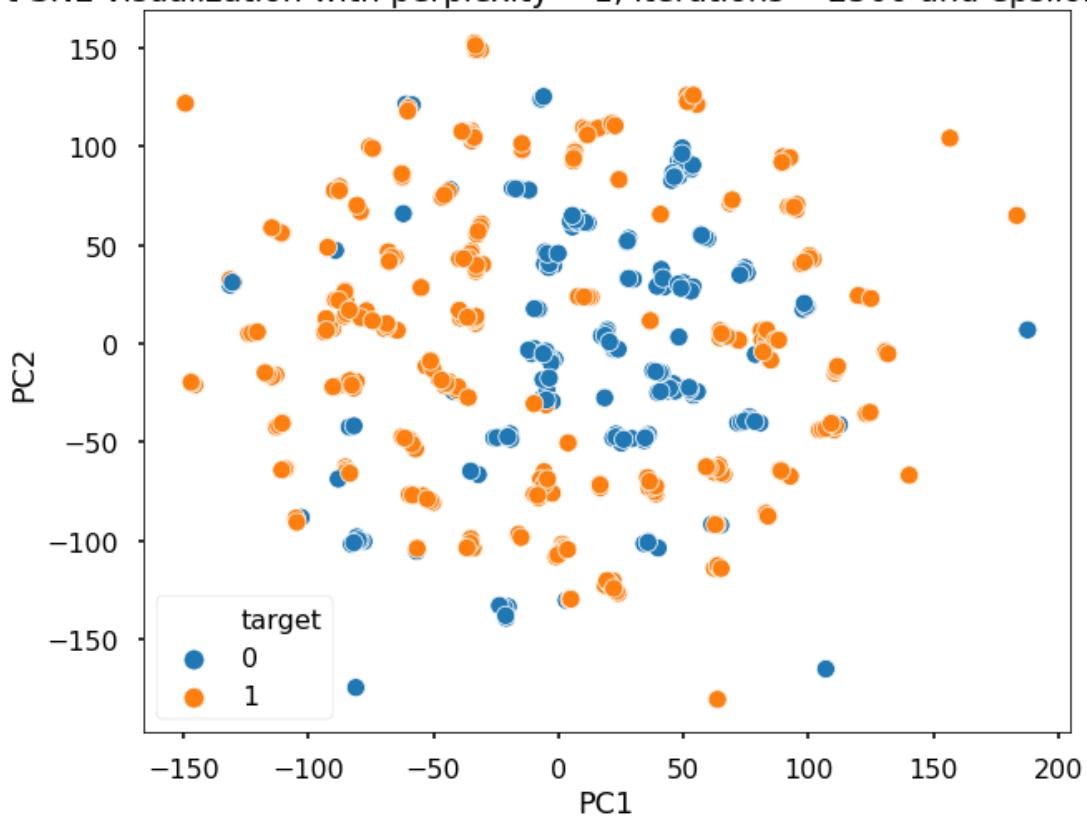
t-SNE visualization with perplexity -- 1, iterations -- 750 and epsilon -- 200



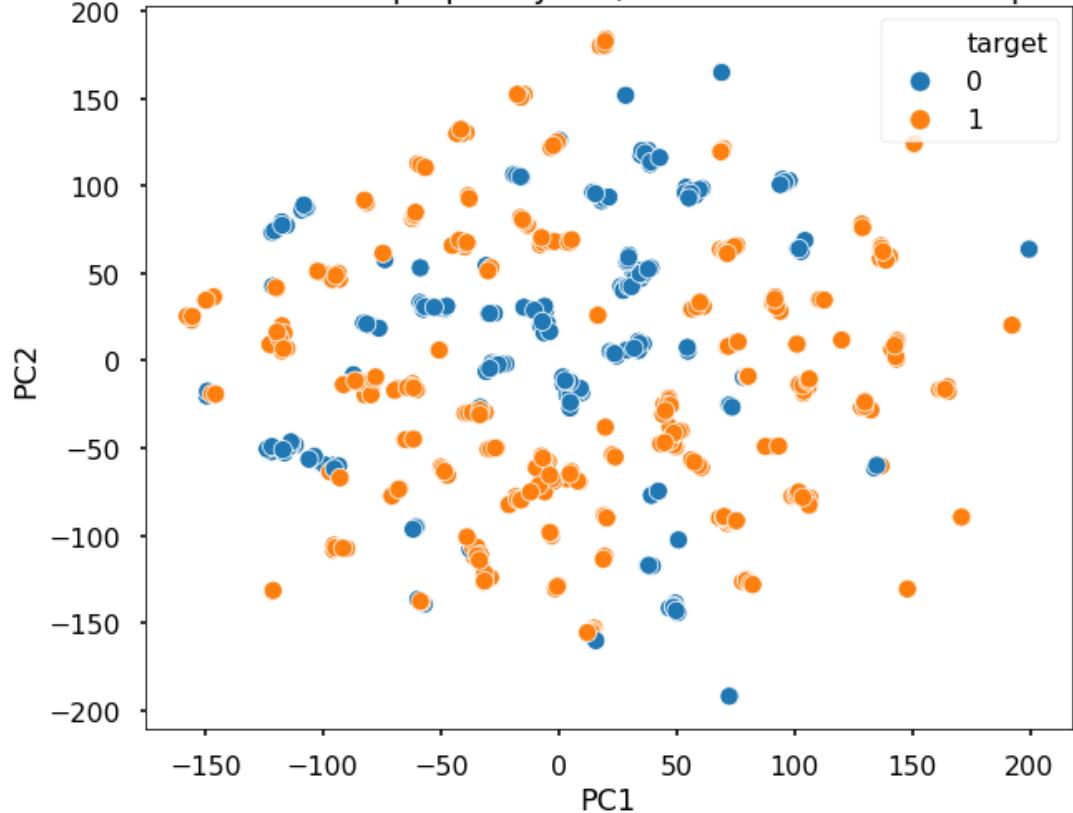
t-SNE visualization with perplexity -- 1, iterations -- 1000 and epsilon -- 200



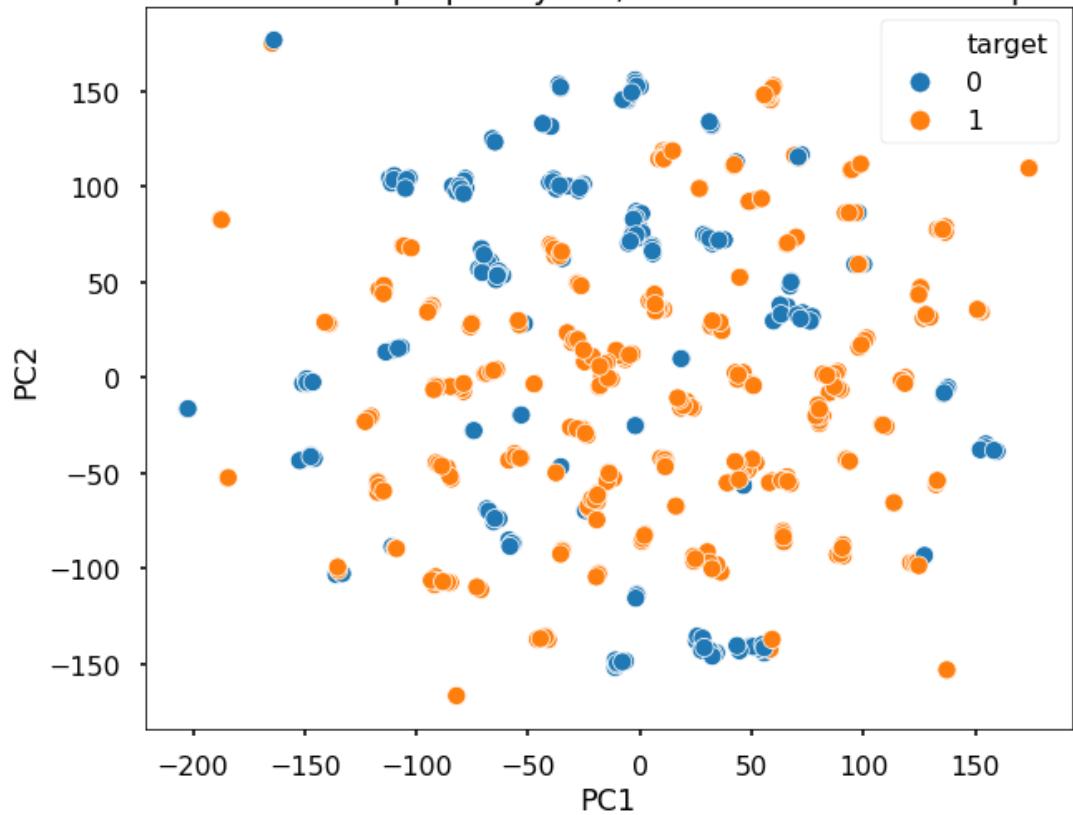
t-SNE visualization with perplexity -- 1, iterations -- 2500 and epsilon -- 200



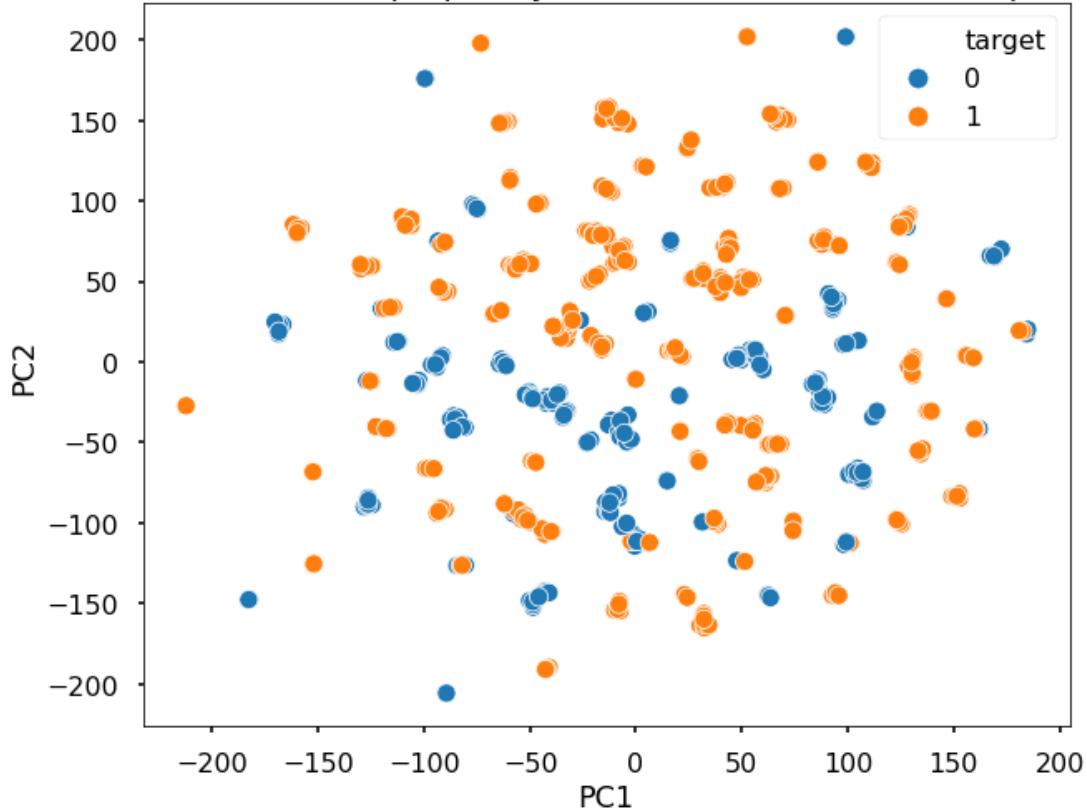
t-SNE visualization with perplexity -- 1, iterations -- 3500 and epsilon -- 200



t-SNE visualization with perplexity -- 1, iterations -- 4000 and epsilon -- 200



t-SNE visualization with perplexity -- 1, iterations -- 5000 and epsilon -- 200



With the lowest value of Perplexity as 1 and Learning rate(Epsilon) slightly on a higher side as 200, we got the small random clusters of both the classes.

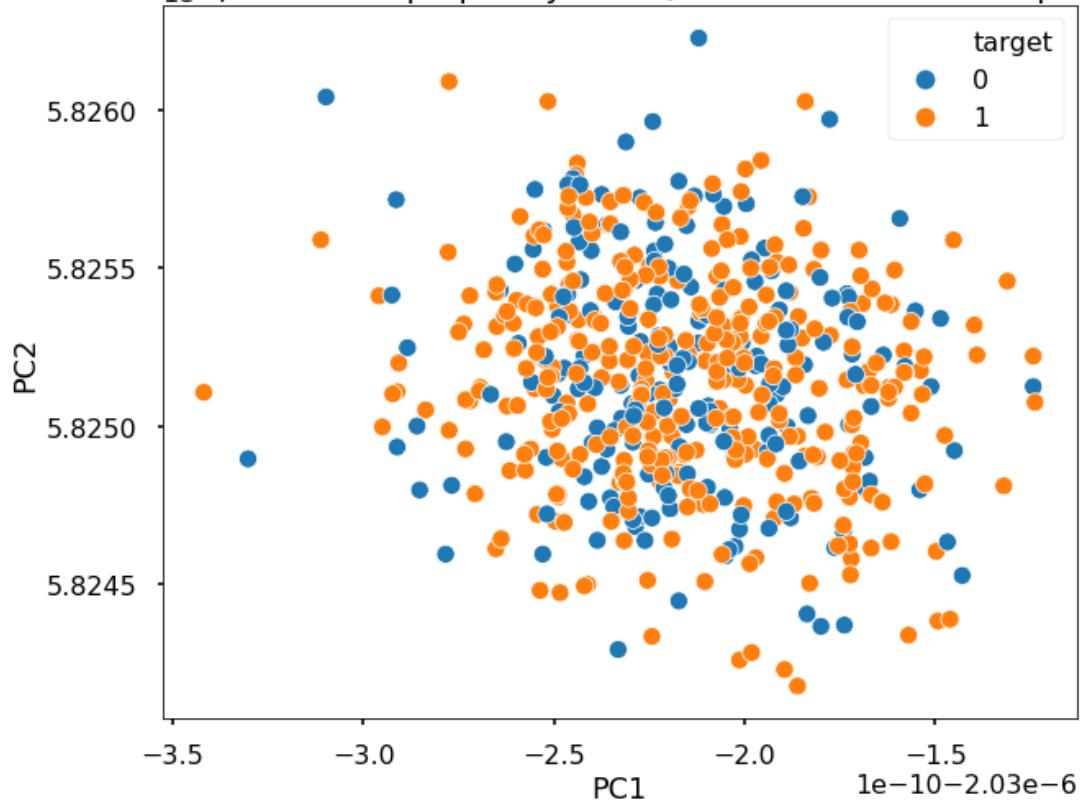
Boundary: CASE-II

- *Highest boundary value of Perplexity as 569(number of data points) with Epsilon as Low as 20 on different Iterations*

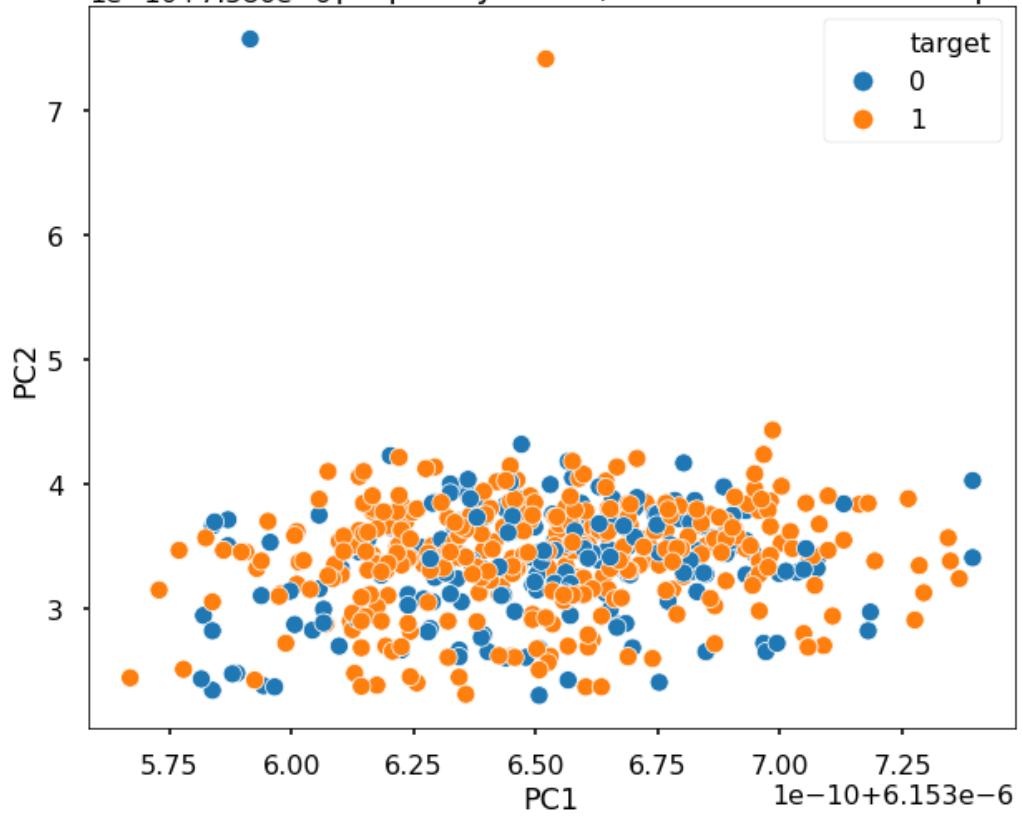
```
In [5]: iterations = [250,350,500,750,1000,2500,3500,4000,5000]

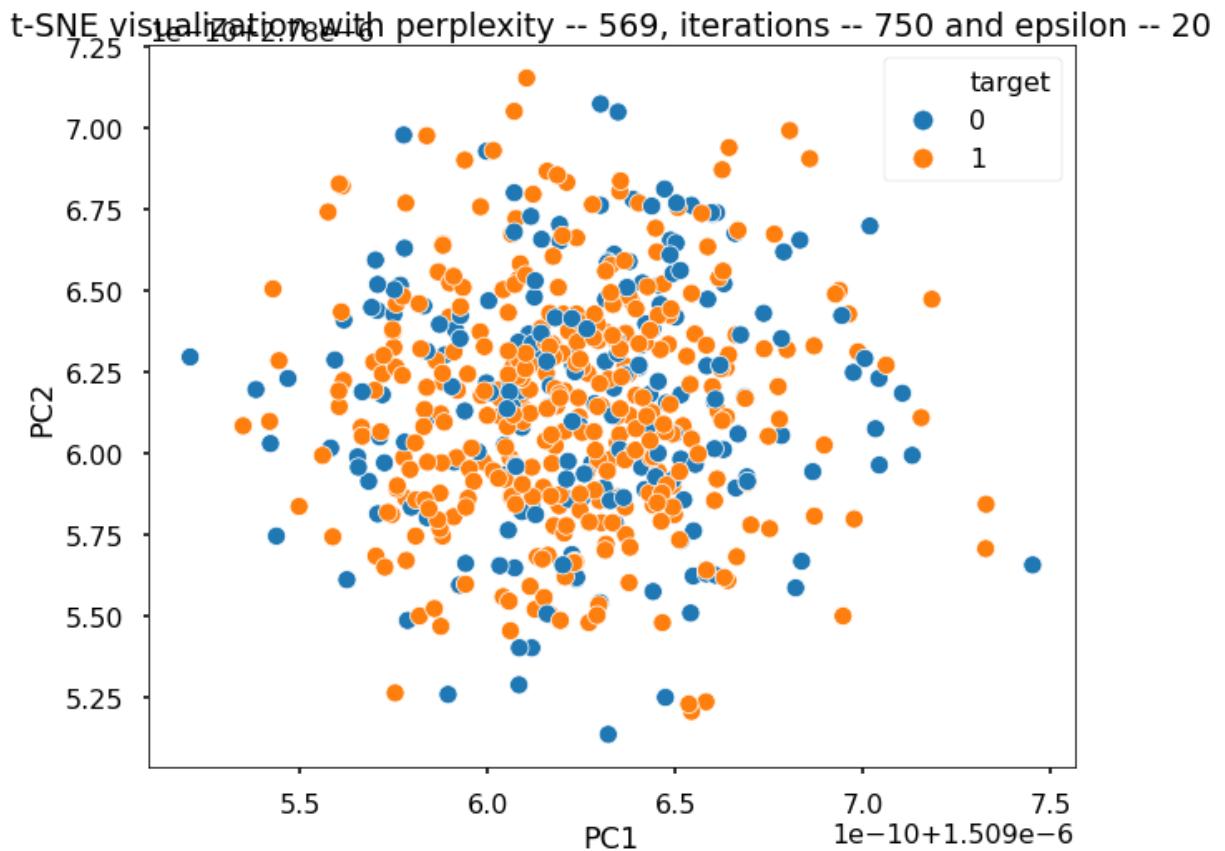
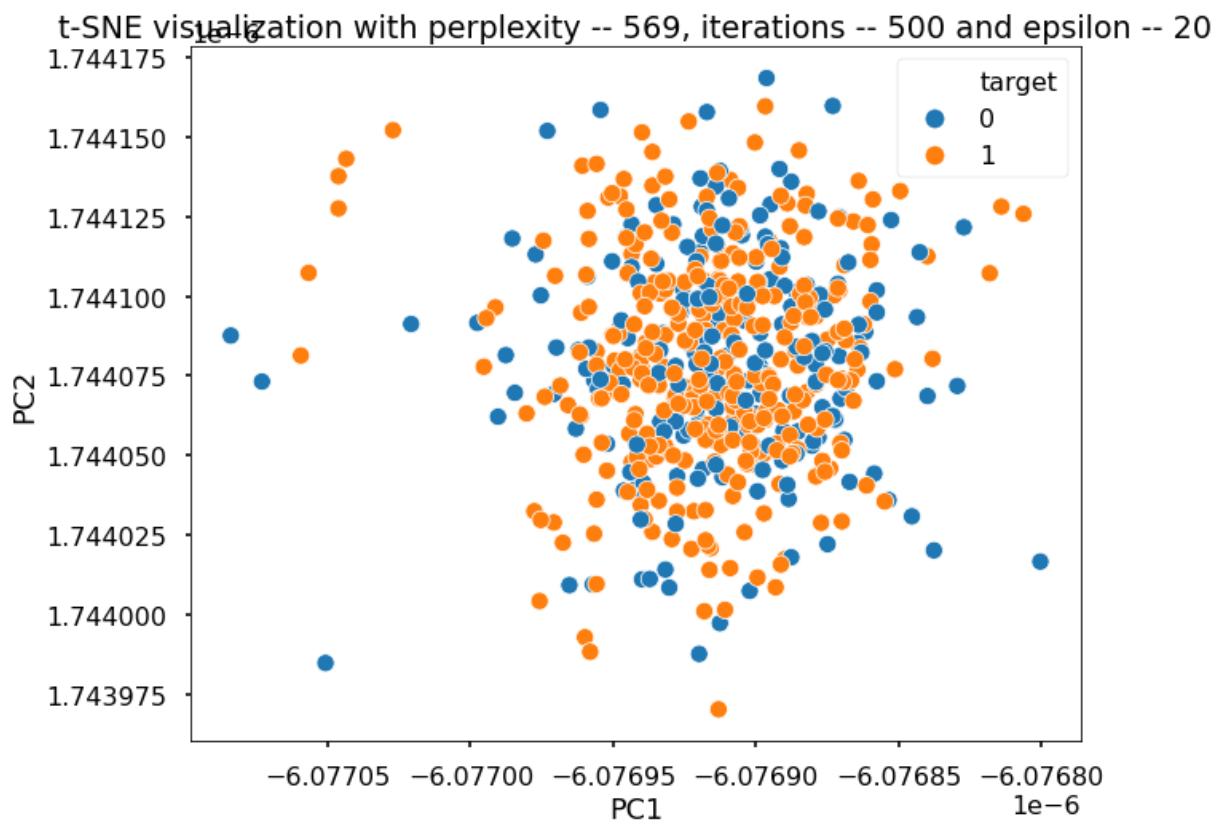
for idx in range(len(iterations)):
    tsne1 = TSNE(n_components=2,perplexity=569,learning_rate=20,n_iter=iterations[idx])
    cancer_tsne_pcmps = pd.DataFrame(tsne1.fit_transform(cancer_norm_df),columns=['PC1','PC2'])
    cancer_tsne_pcmps = pd.concat([cancer_tsne_pcmps,cancer_df['target']],axis=1)
    with plt.style.context('seaborn-poster'):
        plt.figure(figsize=(10,8))
        sns.scatterplot(data=cancer_tsne_pcmps,x='PC1',y='PC2',hue='target')
        plt.title("t-SNE visualization with perplexity -- {0}, iterations -- {1} and epsilon -- 20".format(perplexity,iterations[idx]))
    plt.show()
```

t-SNE visualization with perplexity -- 569, iterations -- 250 and epsilon -- 20

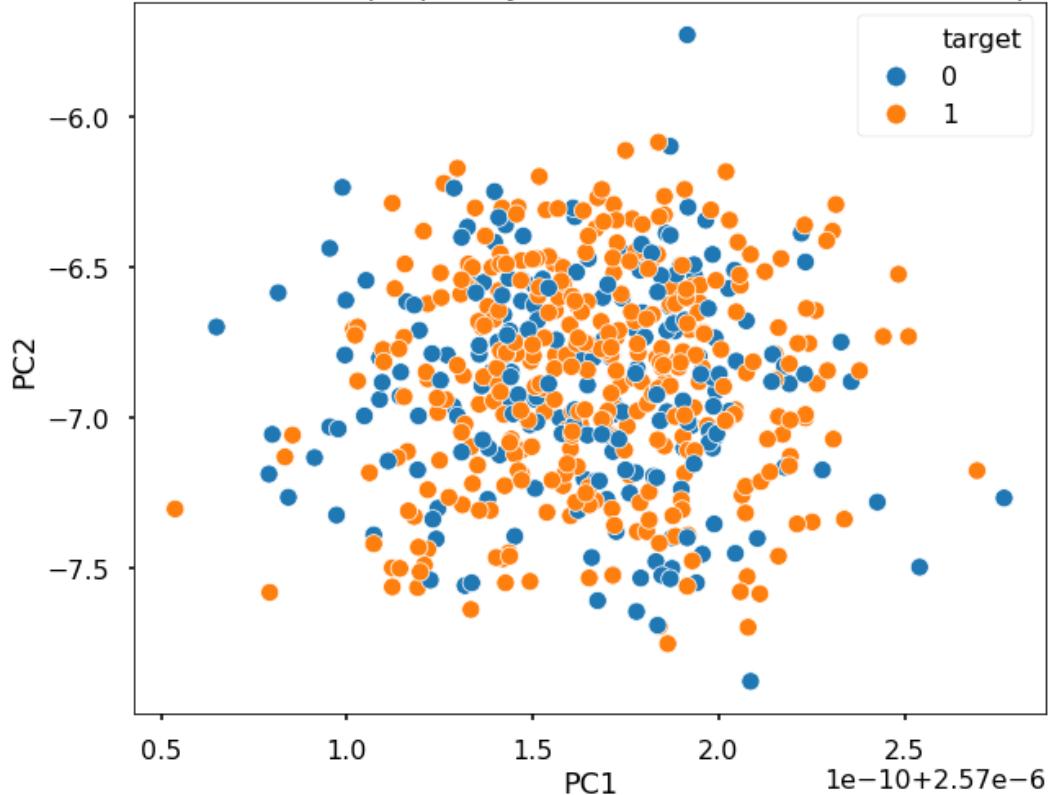


t-SNE visualization with perplexity -- 569, iterations -- 350 and epsilon -- 20

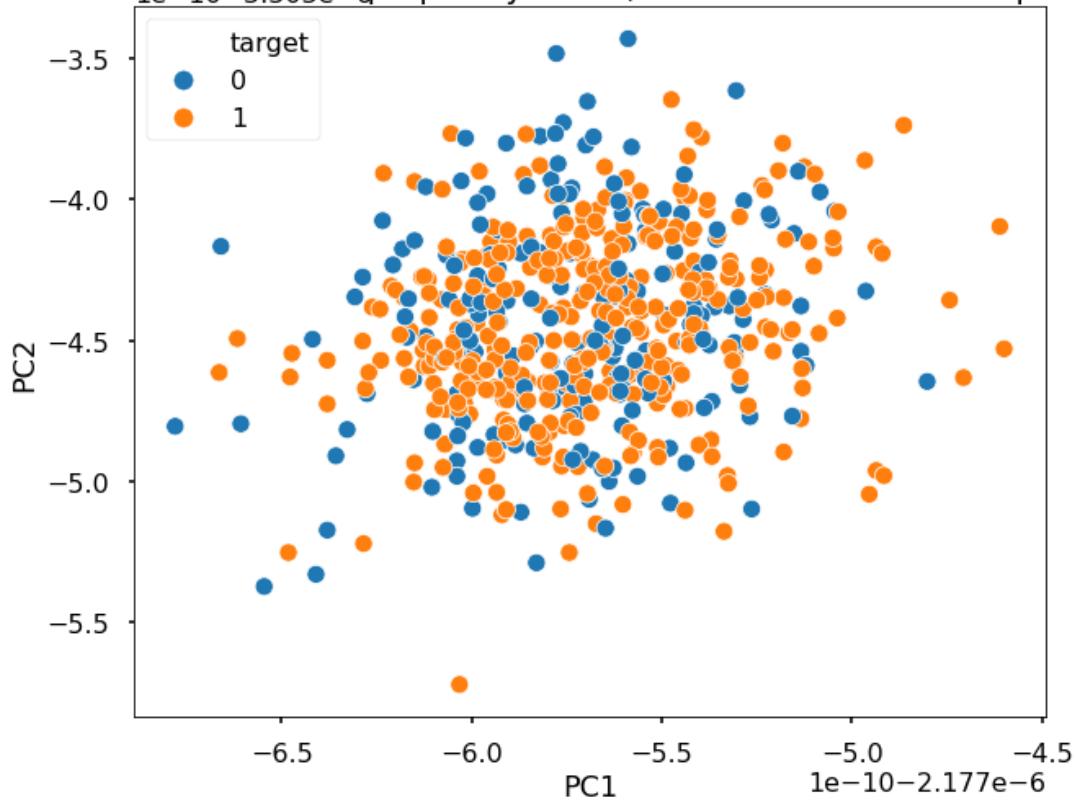




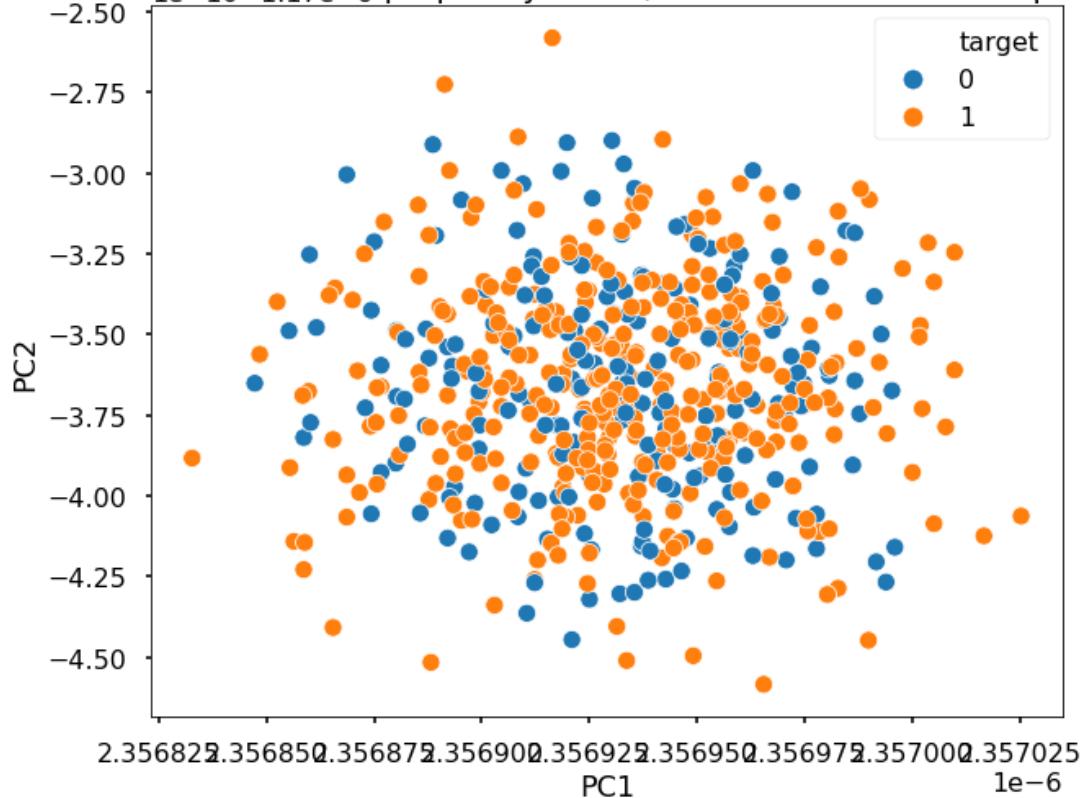
t-SNE visualization with perplexity -- 569, iterations -- 1000 and epsilon -- 20



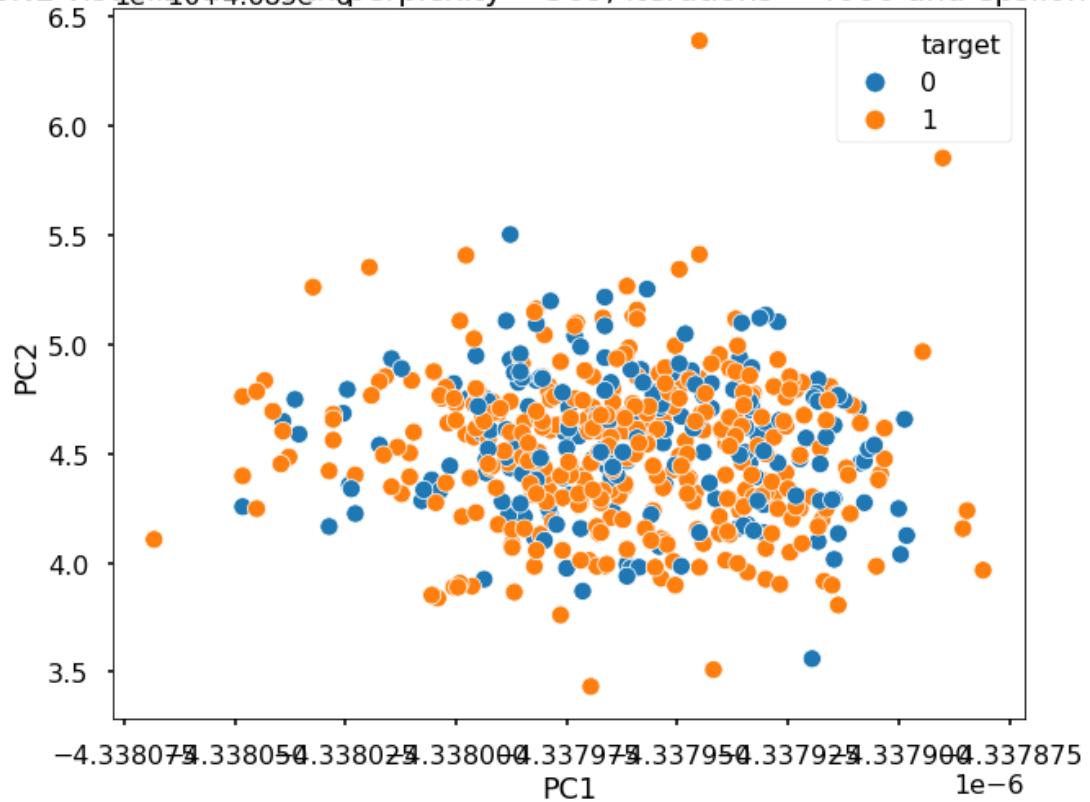
t-SNE visualization with perplexity -- 569, iterations -- 2500 and epsilon -- 20

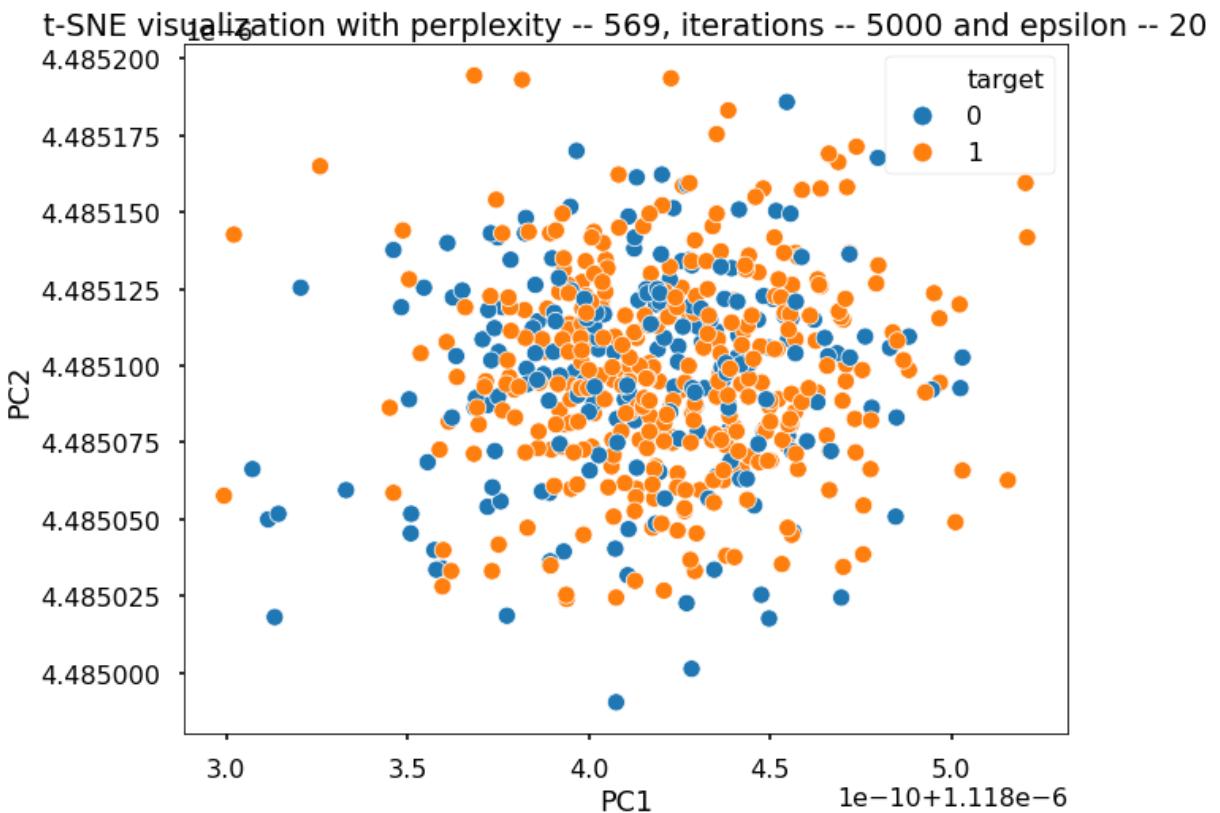


t-SNE visualization with perplexity -- 569, iterations -- 3500 and epsilon -- 20



t-SNE visualization with perplexity -- 569, iterations -- 4000 and epsilon -- 20





With the highest value of Perplexity as 569 and Learning rate(Epsilon) on a lower side as 20, we got the completely random data as a Noise.

Boundary: CASE-III

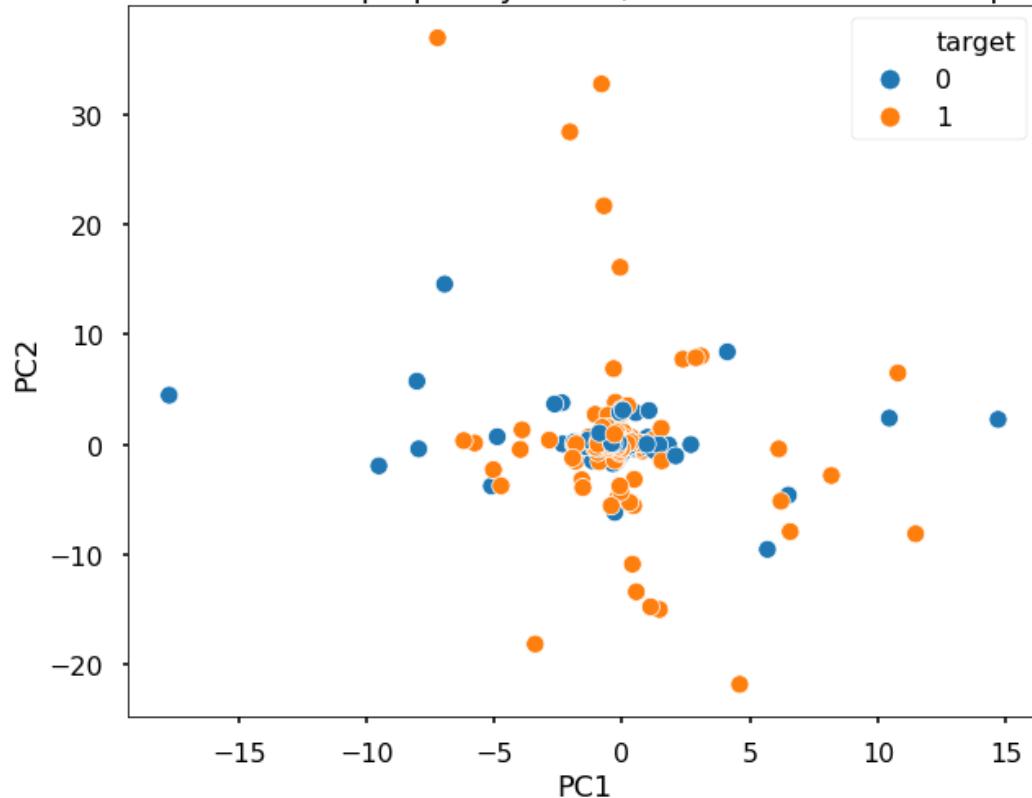
- *Highest boundary value of Perplexity as 569(number of data points) with Epsilon as high as 300 on different Iterations*

```
In [6]: iterations = [250,350,500,750,1000,2500,3500,4000,5000]

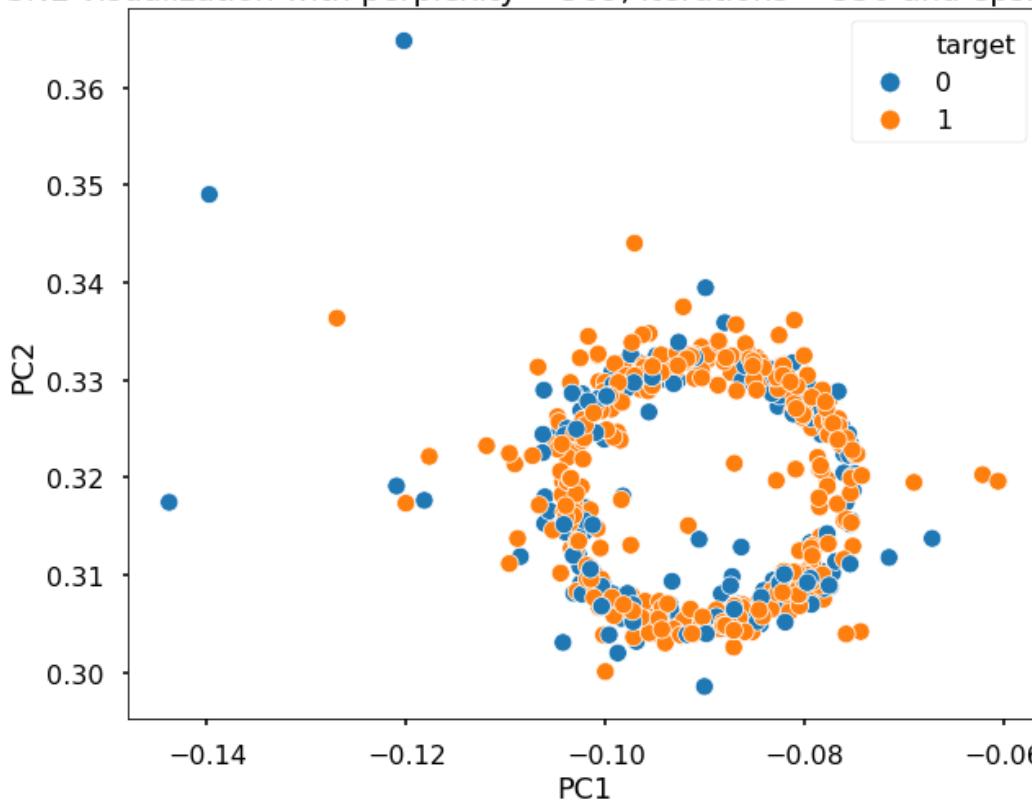
for idx in range(len(iterations)):
    tsne1 = TSNE(n_components=2,perplexity=569,learning_rate=300,n_iter=iterations[i]
    cancer_tsne_pcmps = pd.DataFrame(tsne1.fit_transform(cancer_norm_df),columns=['P
    cancer_tsne_pcmps = pd.concat([cancer_tsne_pcmps,cancer_df['target']],axis=1)
    with plt.style.context('seaborn-poster'):
        plt.figure(figsize=(10,8))
        sns.scatterplot(data=cancer_tsne_pcmps,x='PC1',y='PC2',hue='target')
        plt.title("t-SNE visualization with perplexity -- {0}, iterations -- {1} and

    plt.show()
```

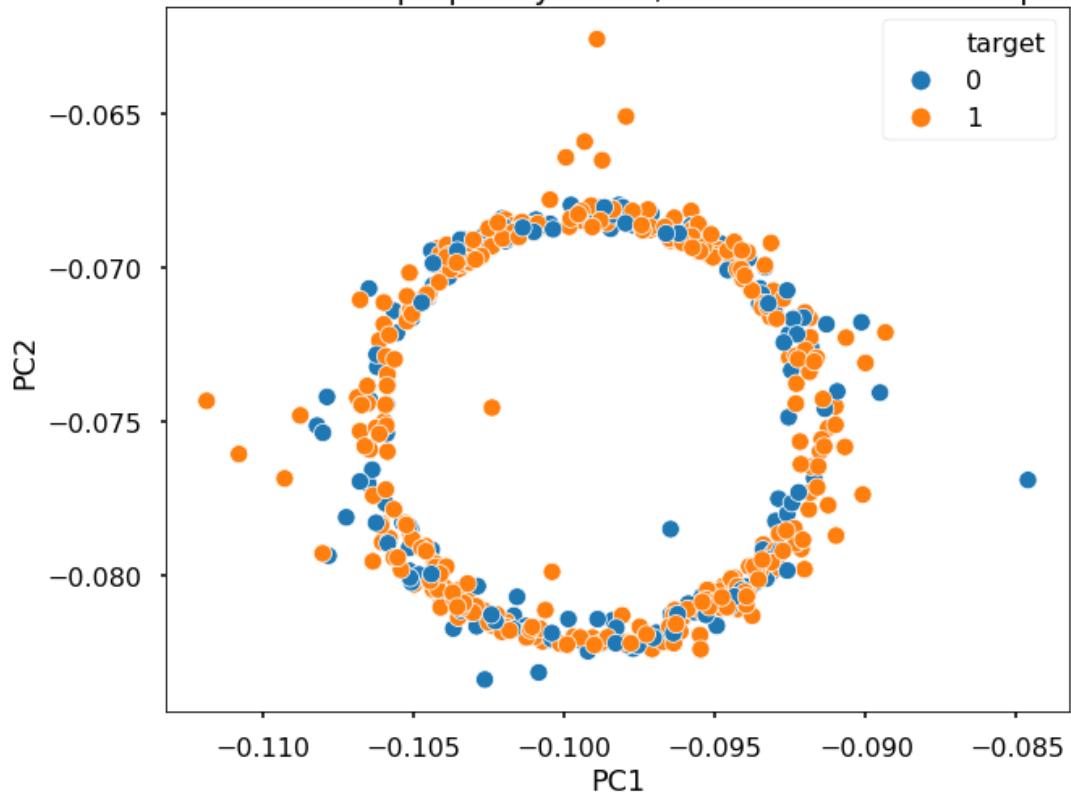
t-SNE visualization with perplexity -- 569, iterations -- 250 and epsilon -- 300



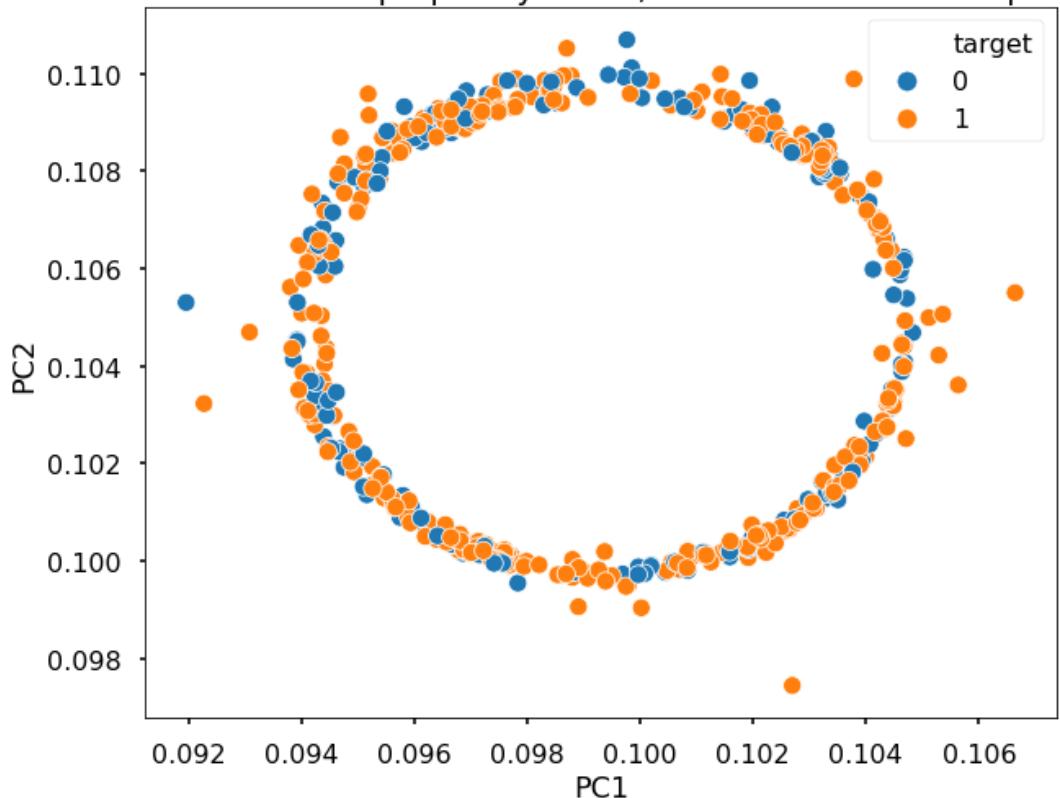
t-SNE visualization with perplexity -- 569, iterations -- 350 and epsilon -- 300



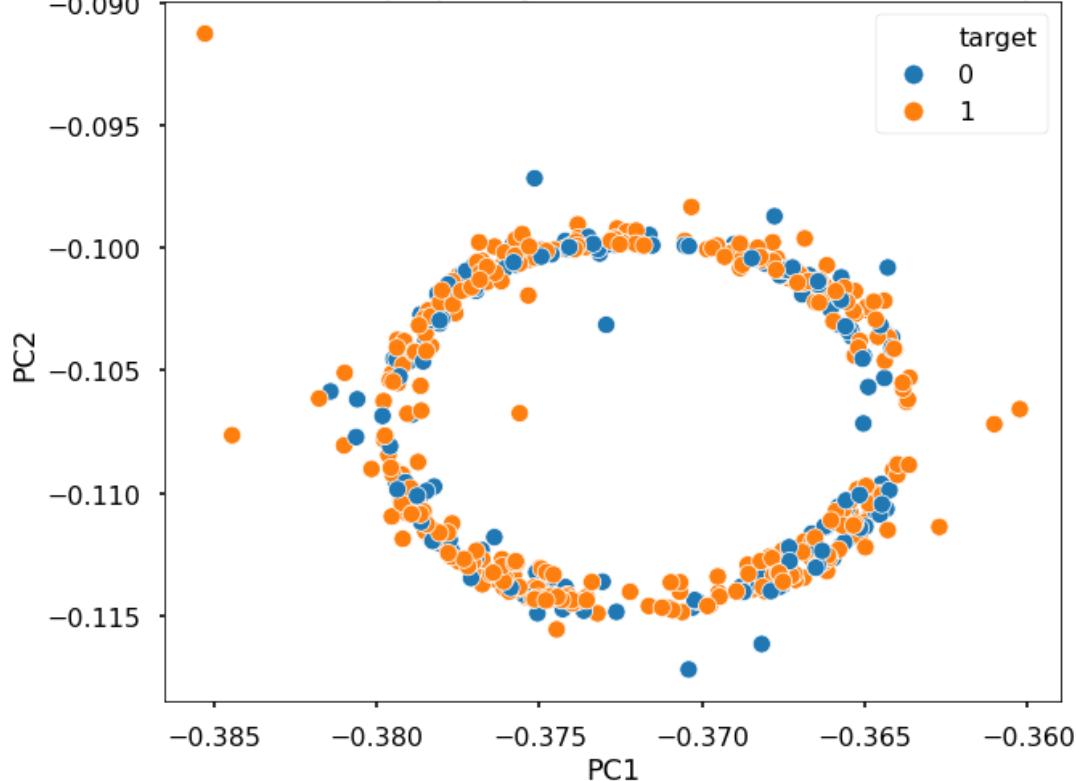
t-SNE visualization with perplexity -- 569, iterations -- 500 and epsilon -- 300



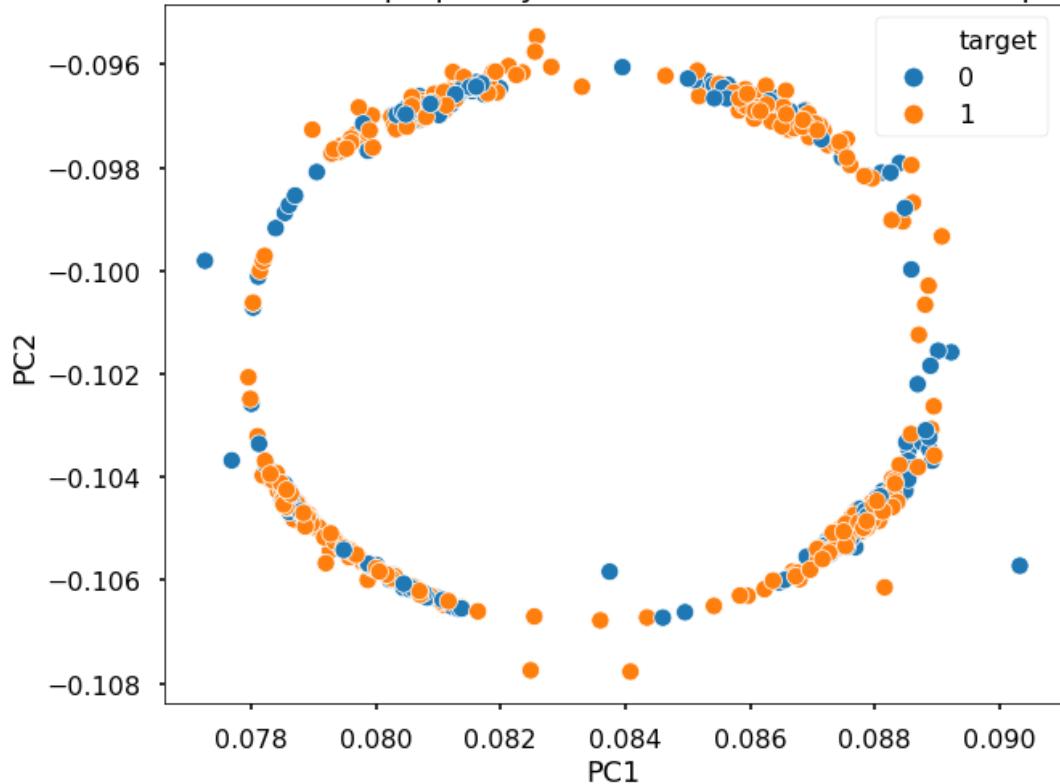
t-SNE visualization with perplexity -- 569, iterations -- 750 and epsilon -- 300



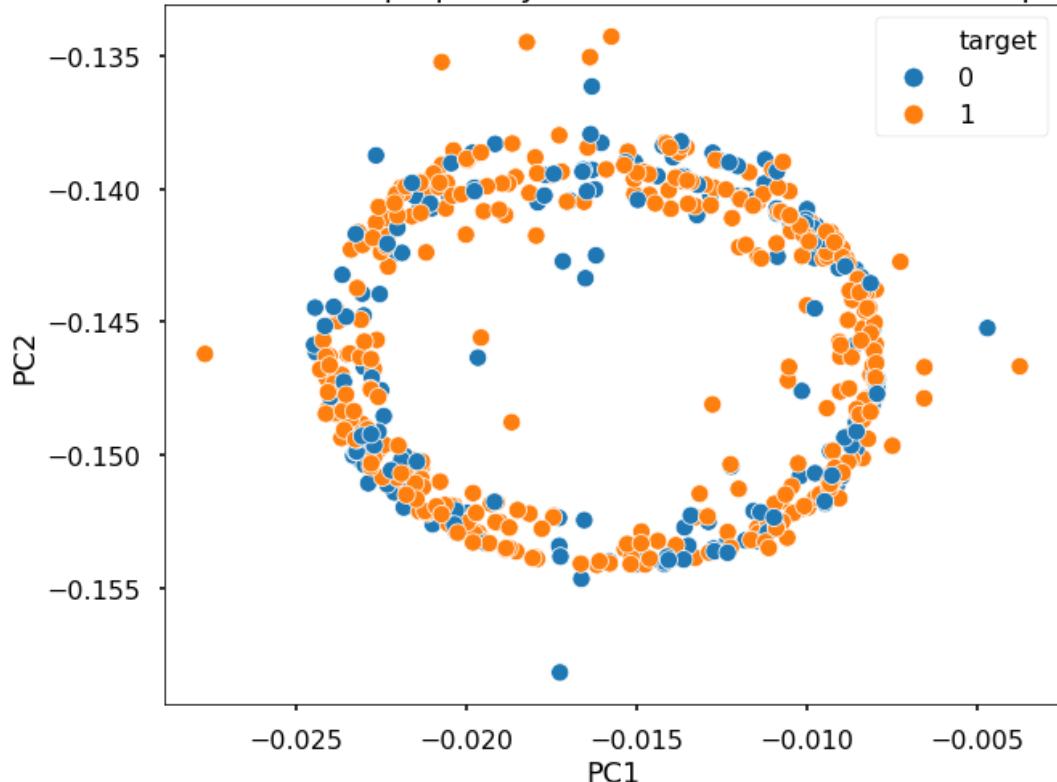
t-SNE visualization with perplexity -- 569, iterations -- 1000 and epsilon -- 300



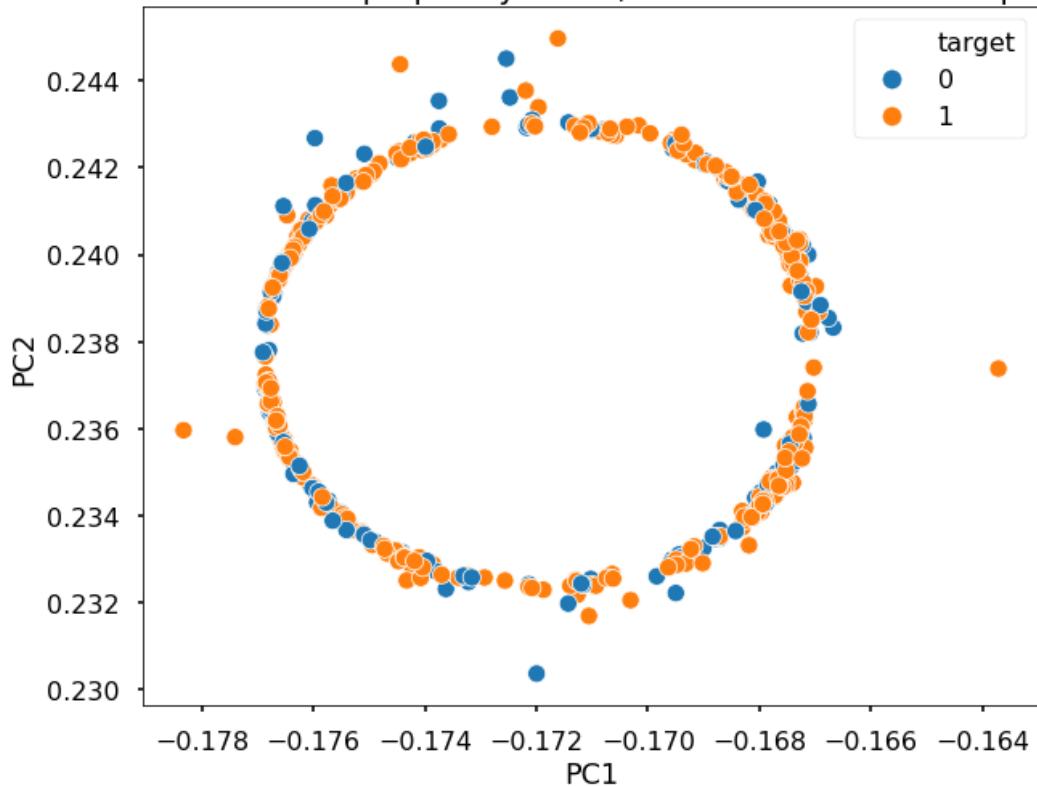
t-SNE visualization with perplexity -- 569, iterations -- 2500 and epsilon -- 300



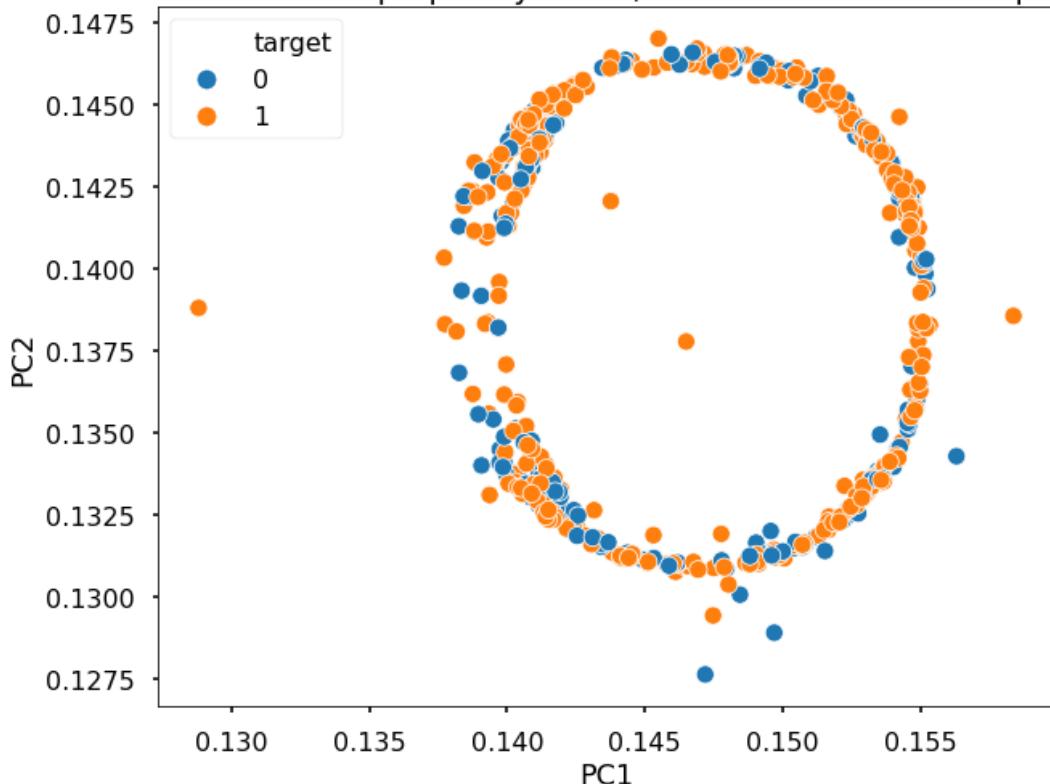
t-SNE visualization with perplexity -- 569, iterations -- 3500 and epsilon -- 300



t-SNE visualization with perplexity -- 569, iterations -- 4000 and epsilon -- 300



t-SNE visualization with perplexity -- 569, iterations -- 5000 and epsilon -- 300



With the highest value of Perplexity as 569 and Learning rate(Epsilon) on a higher side as 300, we got the ball shaped data cluster which says that every point is a definite distance away from each other.

Multiple_runs

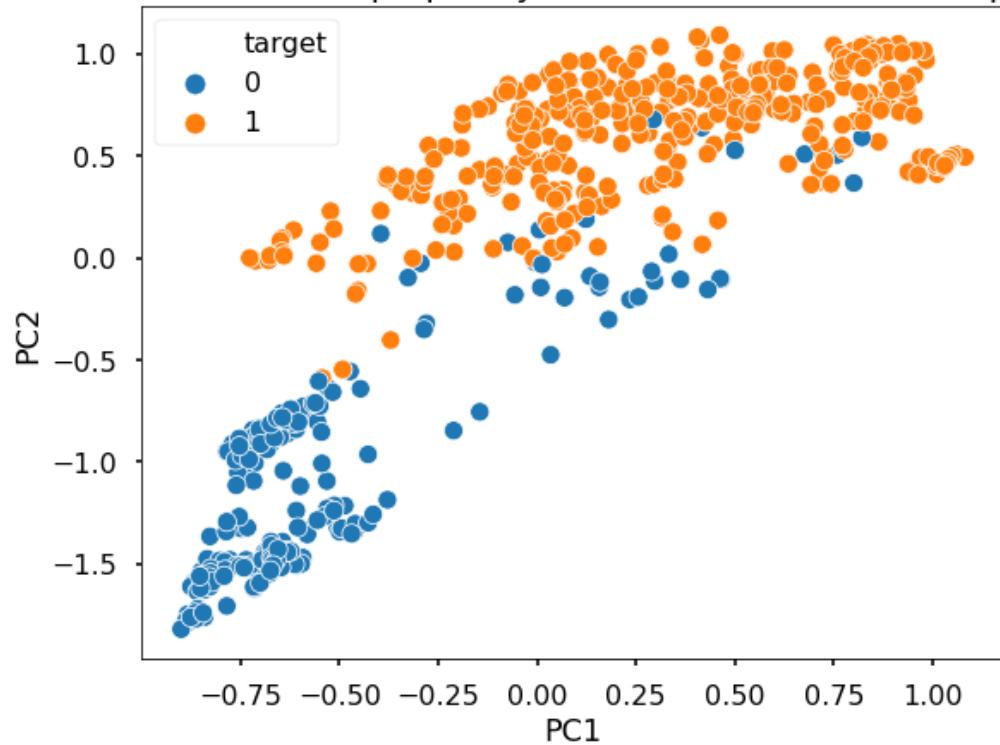
CASE-I-Multiple_runs

Running t-SNE on a range of perplexities, iterations and Learning rates with embedding initialization as 'random'

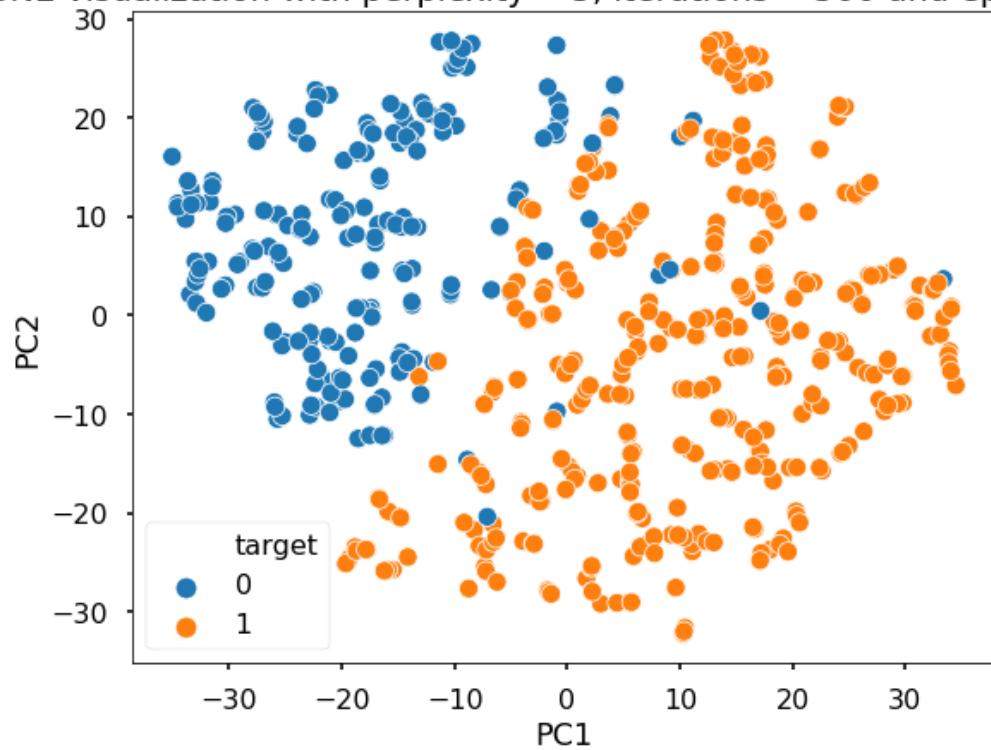
```
In [8]: perplexities = [30,50,100]
iterations = [250,500,750,1000,3000,5000]
l_rate = [10,30,50,100,200]

for pidx in range(len(perplexities)):
    for lidx in range(len(l_rate)):
        for idx in range(len(iterations)):
            tsne1 = TSNE(n_components=2,init='random',perplexity=perplexities[pidx],
                          n_jobs=-1)
            cancer_tsne_pcmps = pd.DataFrame(tsne1.fit_transform(cancer_norm_df),columns=cancer_norm_df.columns)
            cancer_tsne_pcmps = pd.concat([cancer_tsne_pcmps,cancer_df[['target']]],axis=1)
            with plt.style.context('seaborn-poster'):
                plt.figure(figsize=(9,7))
                sns.scatterplot(data=cancer_tsne_pcmps,x='PC1',y='PC2',hue='target')
                plt.title("t-SNE visualization with perplexity -- {0}, iterations -- {1}, learning rate -- {2} ".format(perplexities[pidx],iterations[idx],l_rate[lidx]))
            plt.show()
```

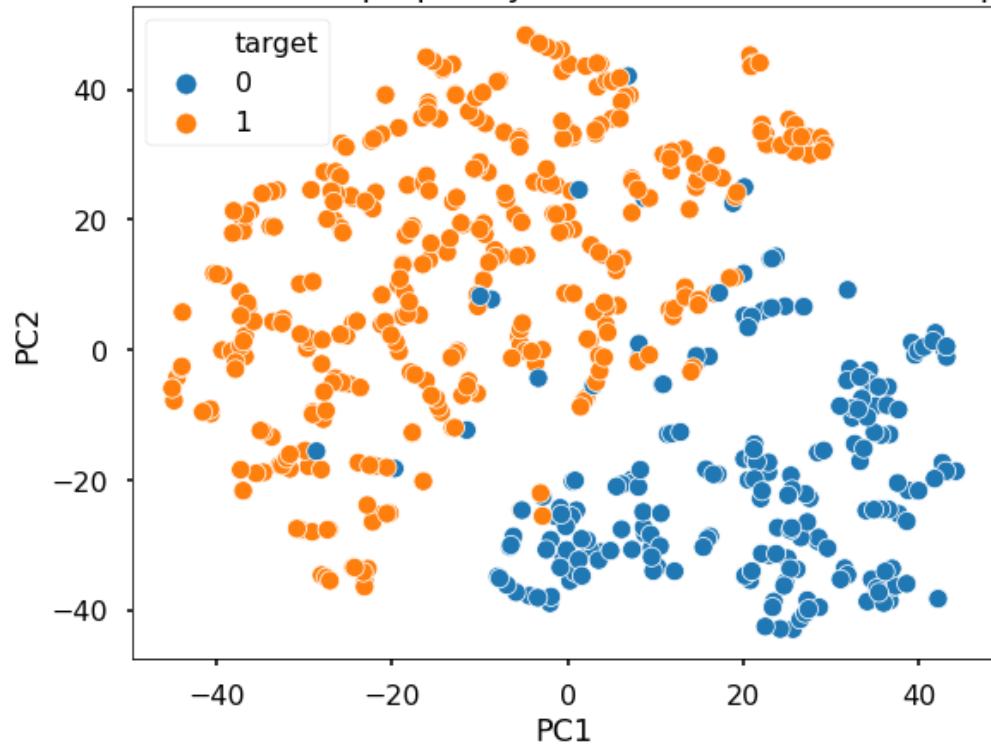
t-SNE visualization with perplexity -- 5, iterations -- 250 and epsilon -- 10



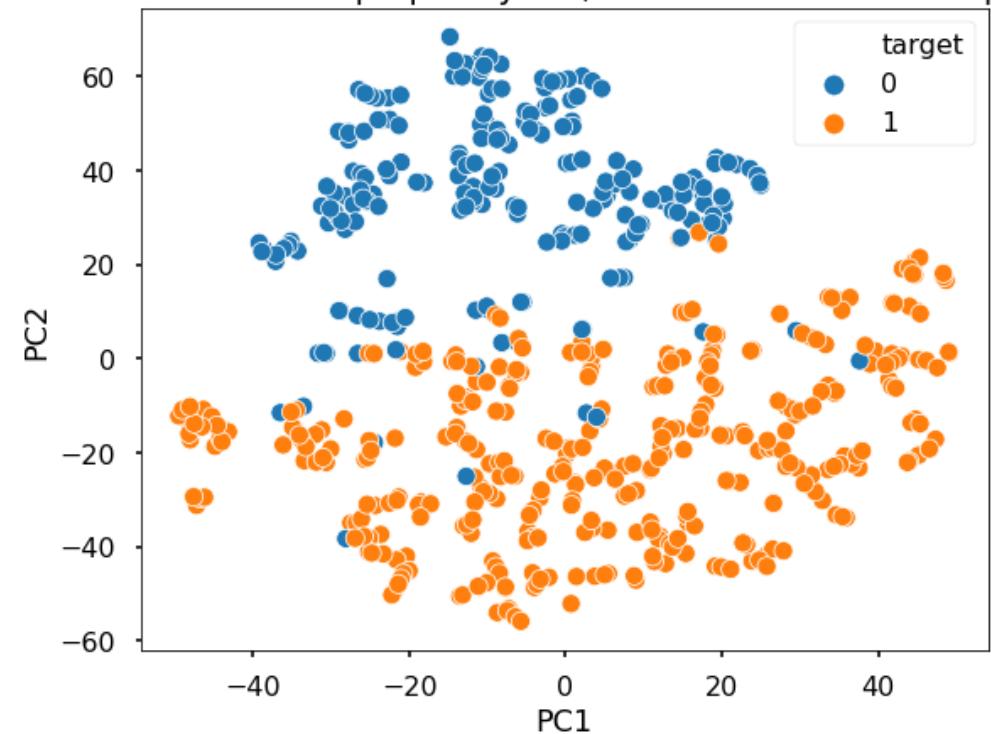
t-SNE visualization with perplexity -- 5, iterations -- 500 and epsilon -- 10



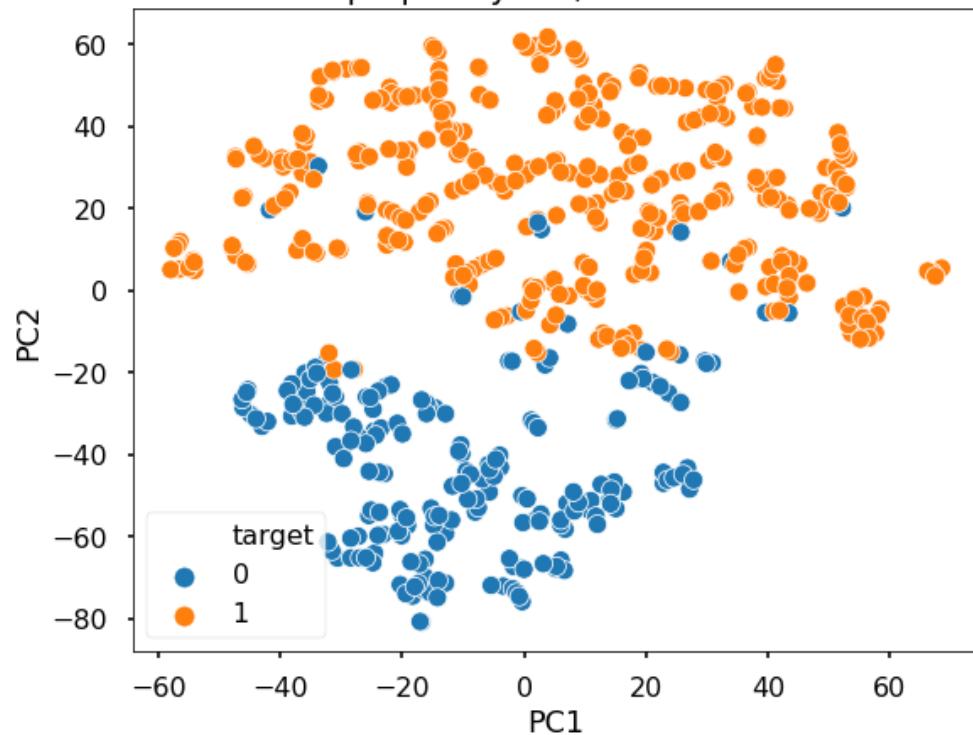
t-SNE visualization with perplexity -- 5, iterations -- 750 and epsilon -- 10



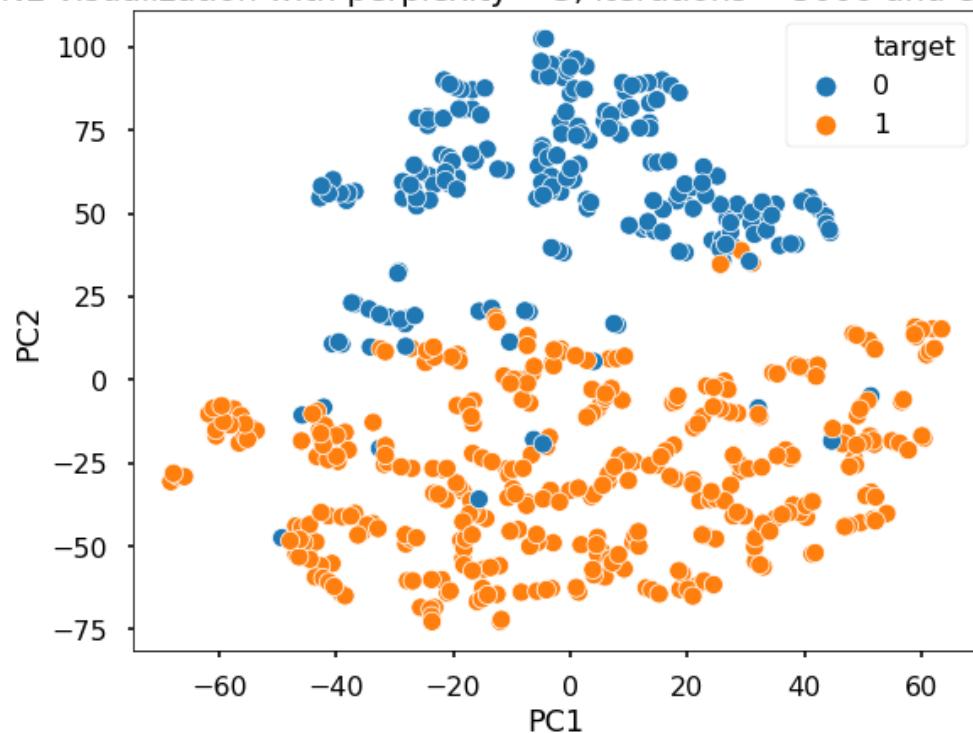
t-SNE visualization with perplexity -- 5, iterations -- 1000 and epsilon -- 10



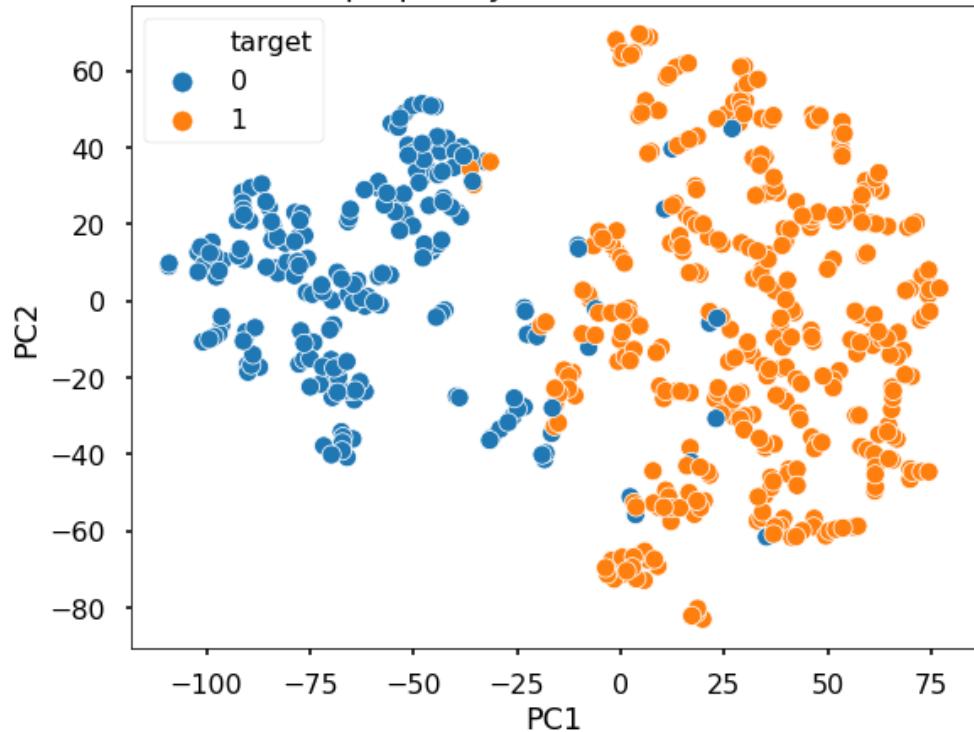
t-SNE visualization with perplexity -- 5, iterations -- 2000 and epsilon -- 10



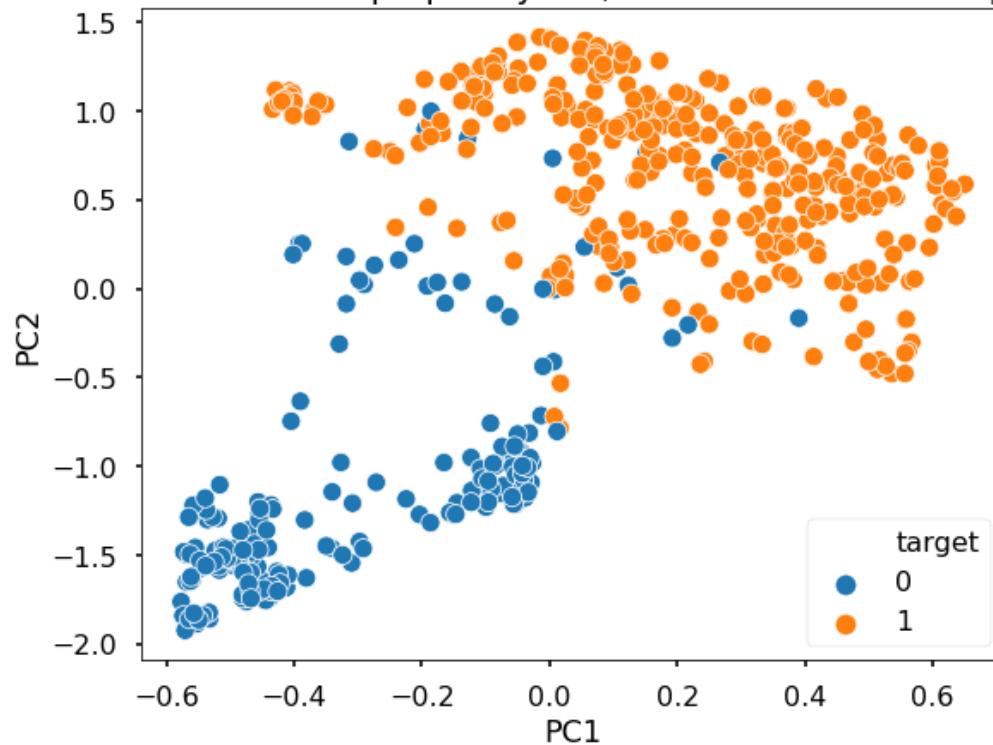
t-SNE visualization with perplexity -- 5, iterations -- 3000 and epsilon -- 10



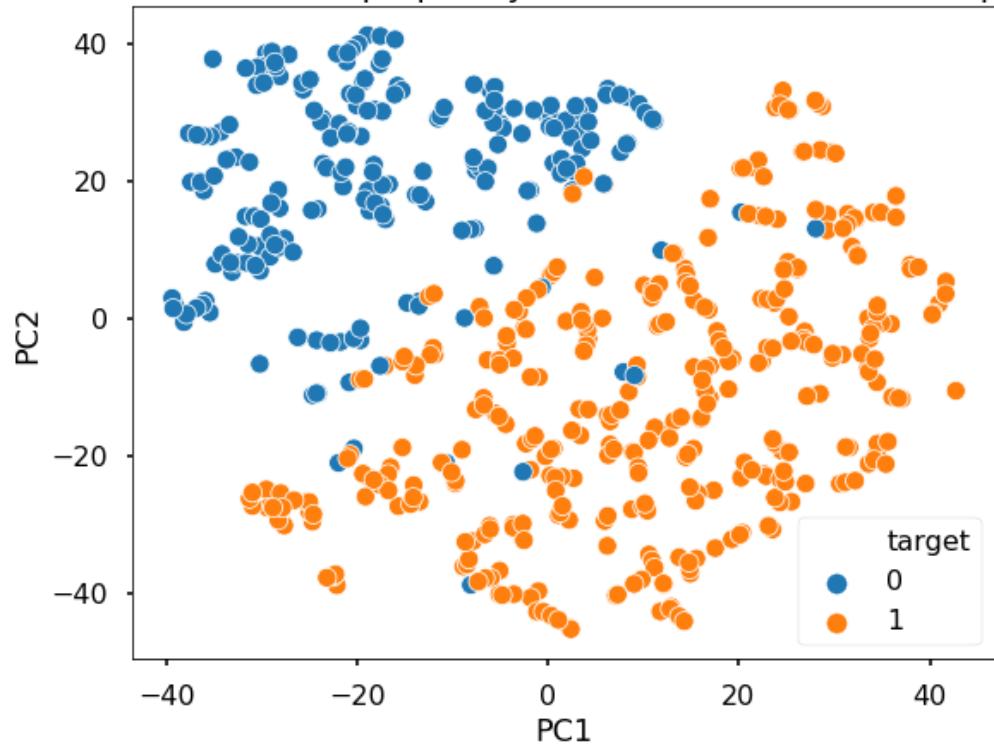
t-SNE visualization with perplexity -- 5, iterations -- 5000 and epsilon -- 10



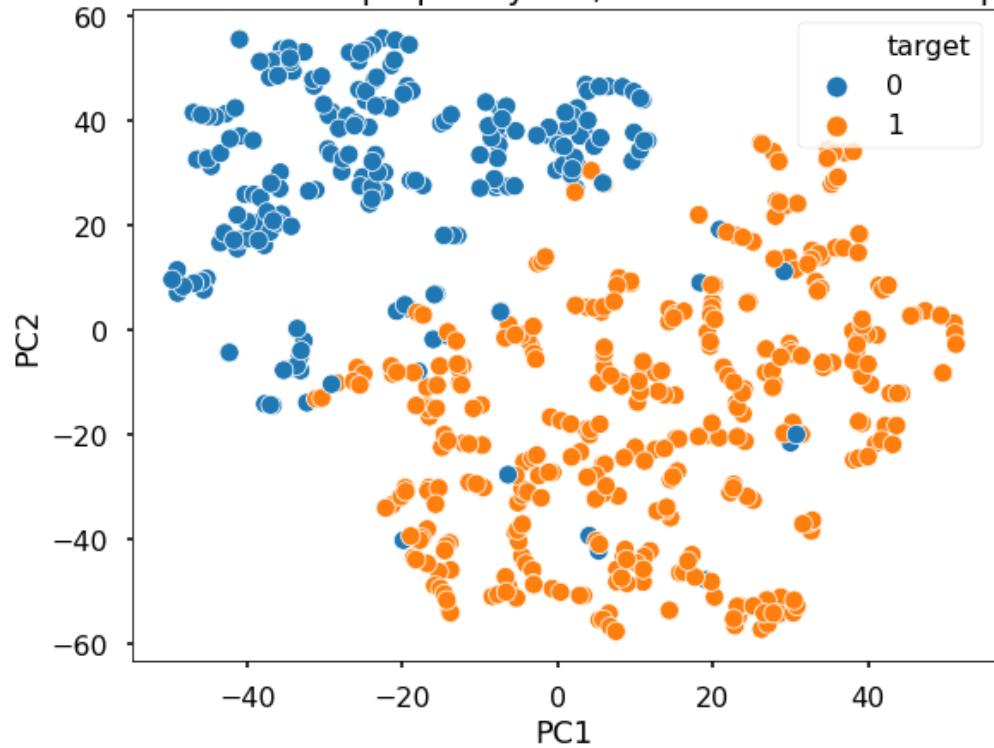
t-SNE visualization with perplexity -- 5, iterations -- 250 and epsilon -- 30



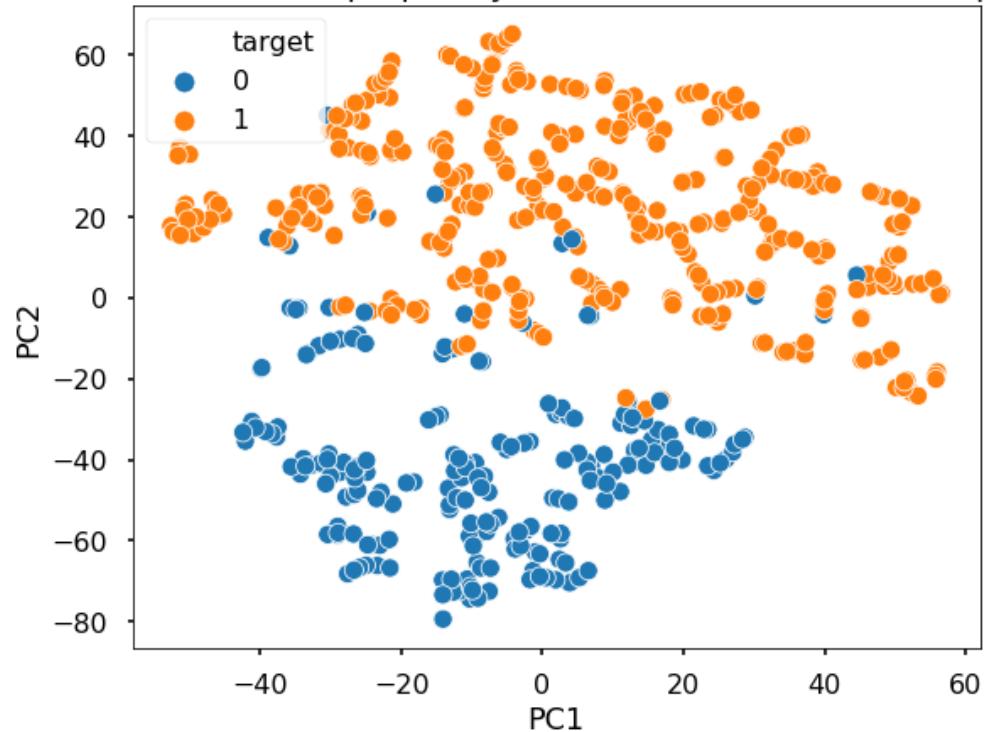
t-SNE visualization with perplexity -- 5, iterations -- 500 and epsilon -- 30



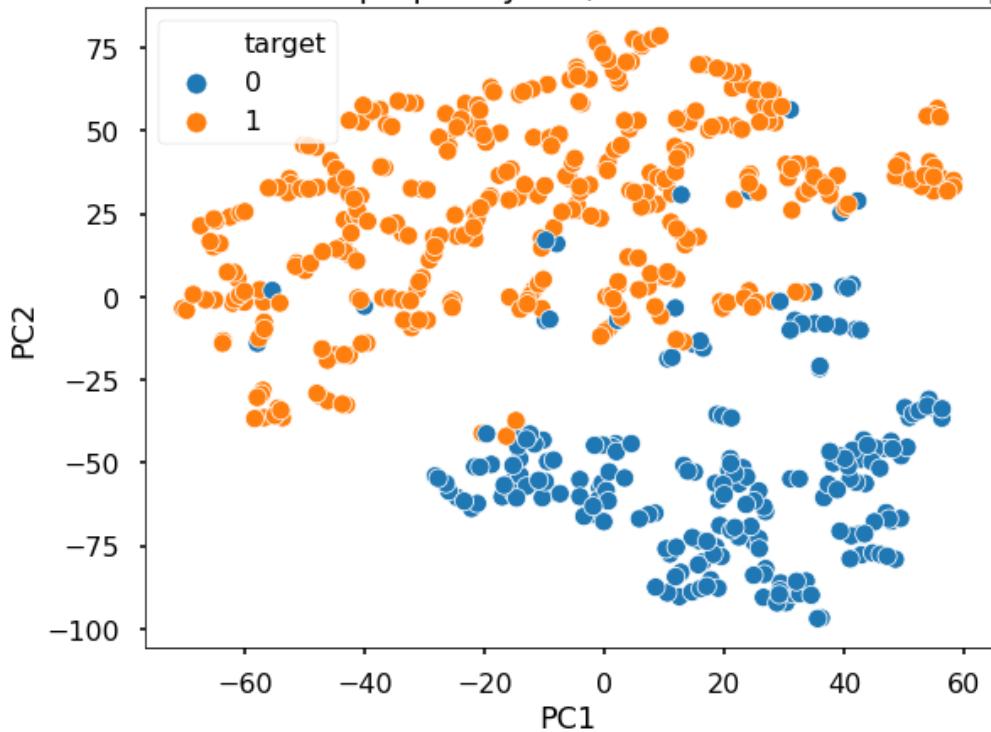
t-SNE visualization with perplexity -- 5, iterations -- 750 and epsilon -- 30



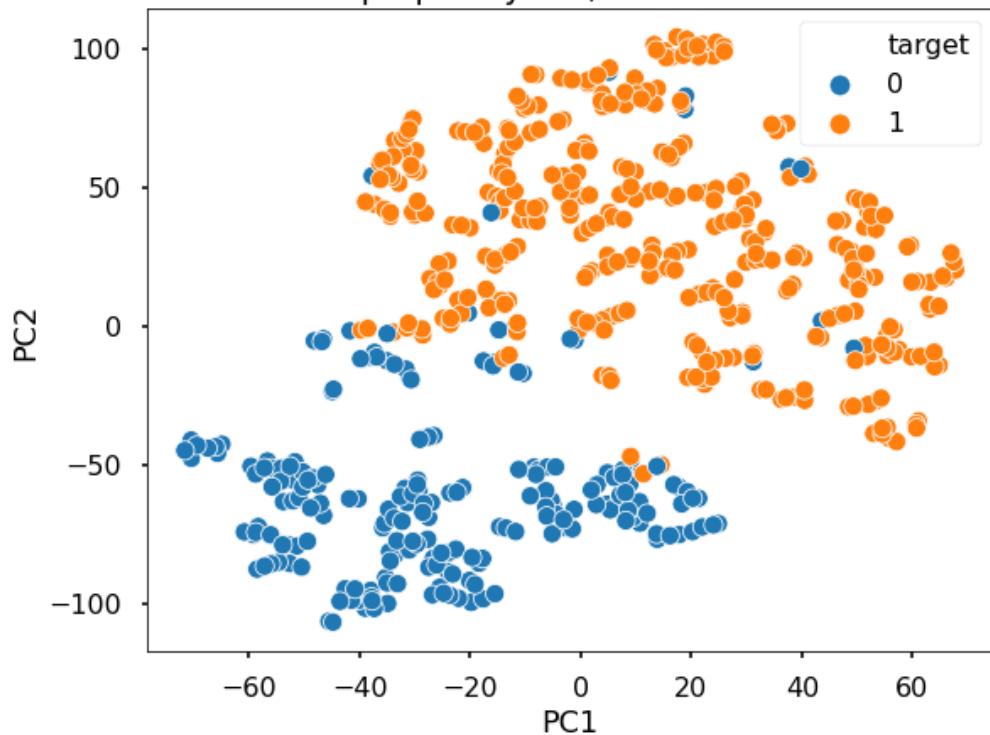
t-SNE visualization with perplexity -- 5, iterations -- 1000 and epsilon -- 30



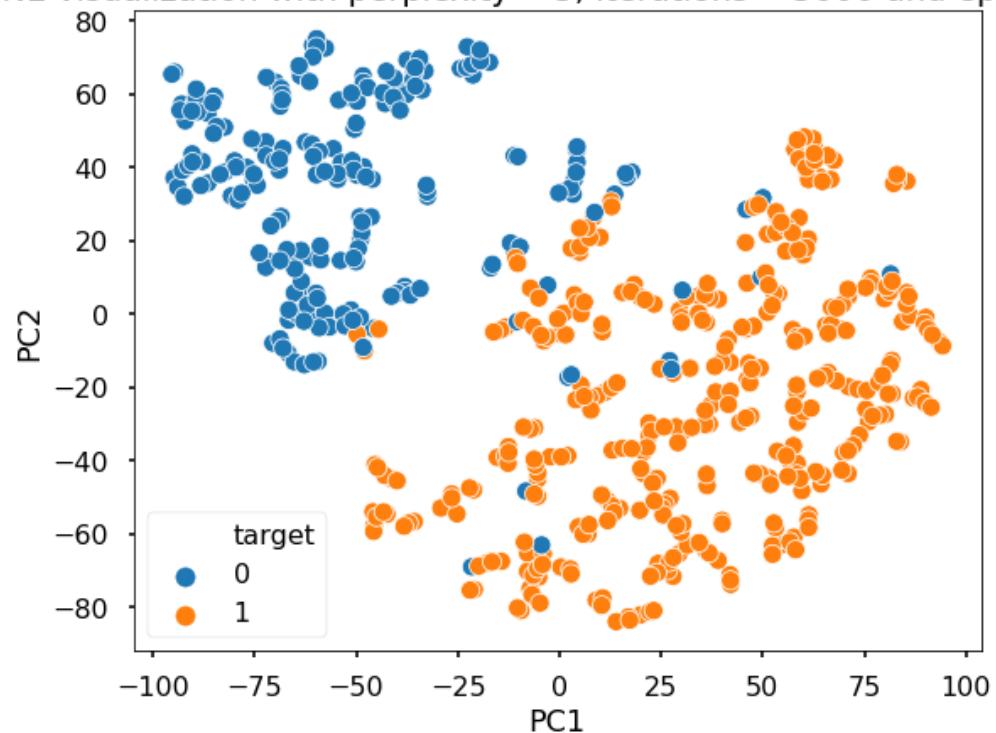
t-SNE visualization with perplexity -- 5, iterations -- 2000 and epsilon -- 30



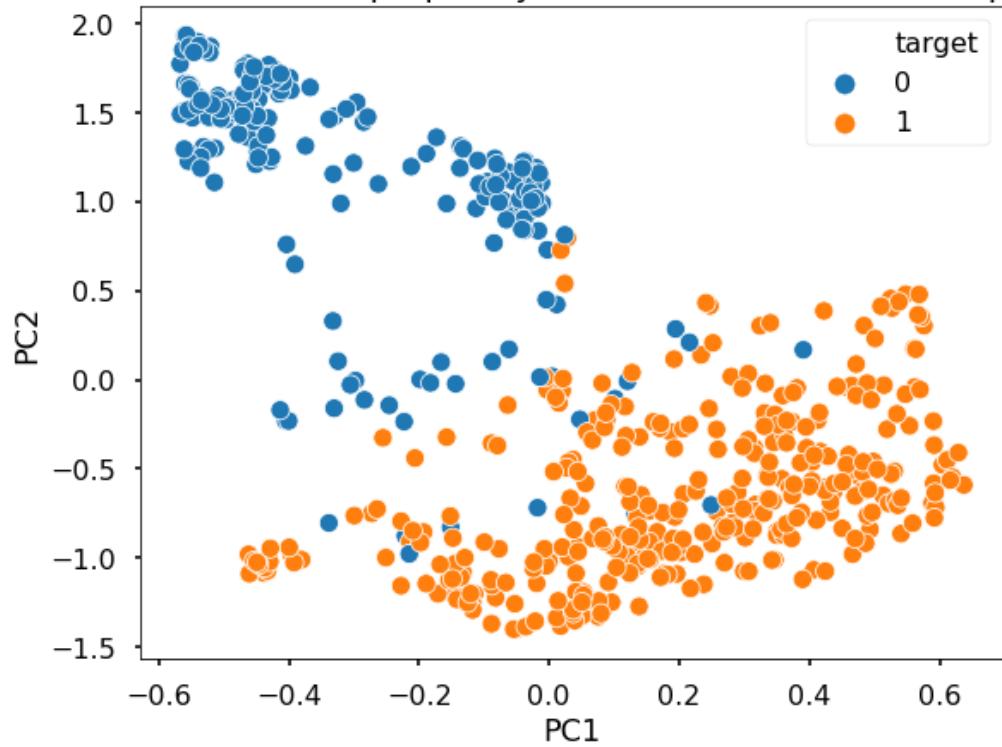
t-SNE visualization with perplexity -- 5, iterations -- 3000 and epsilon -- 30



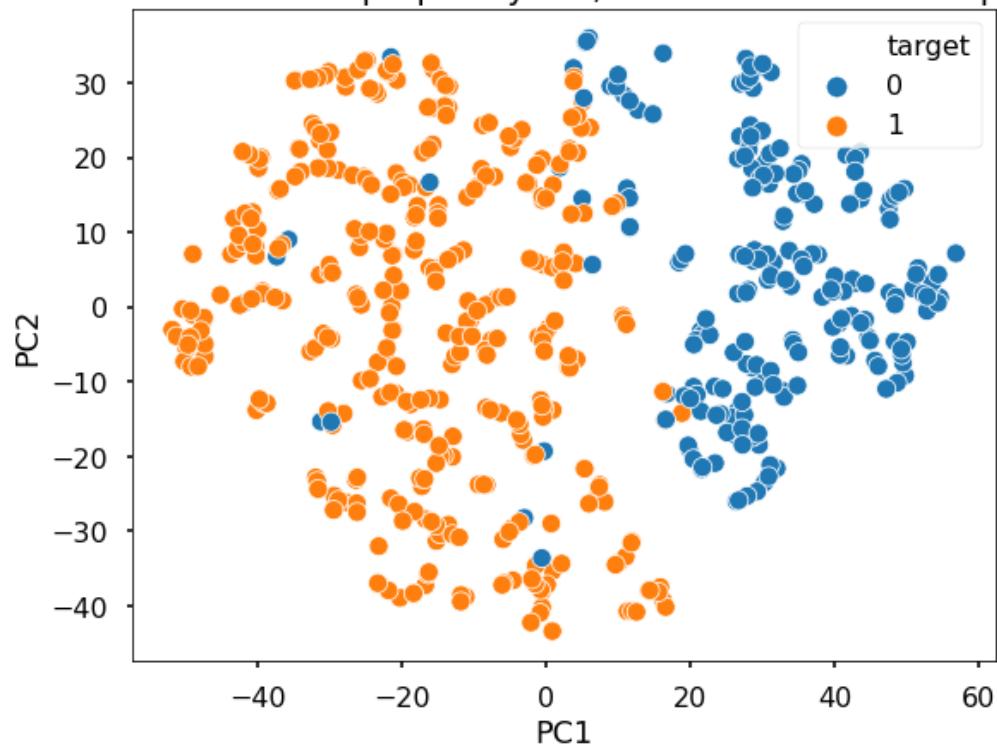
t-SNE visualization with perplexity -- 5, iterations -- 5000 and epsilon -- 30



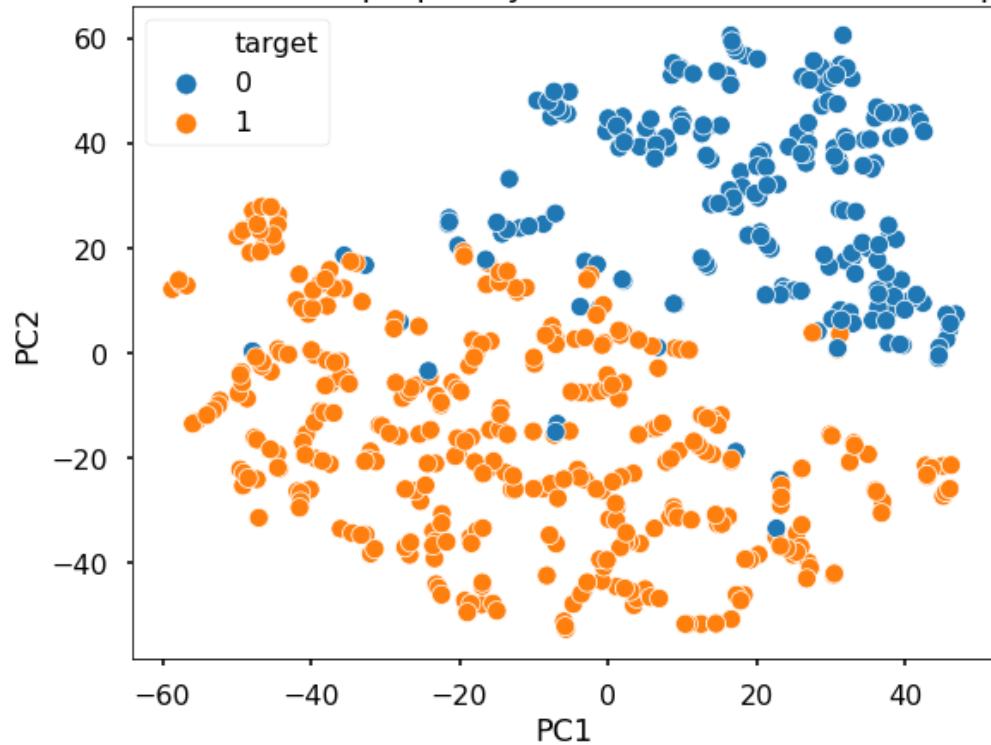
t-SNE visualization with perplexity -- 5, iterations -- 250 and epsilon -- 50



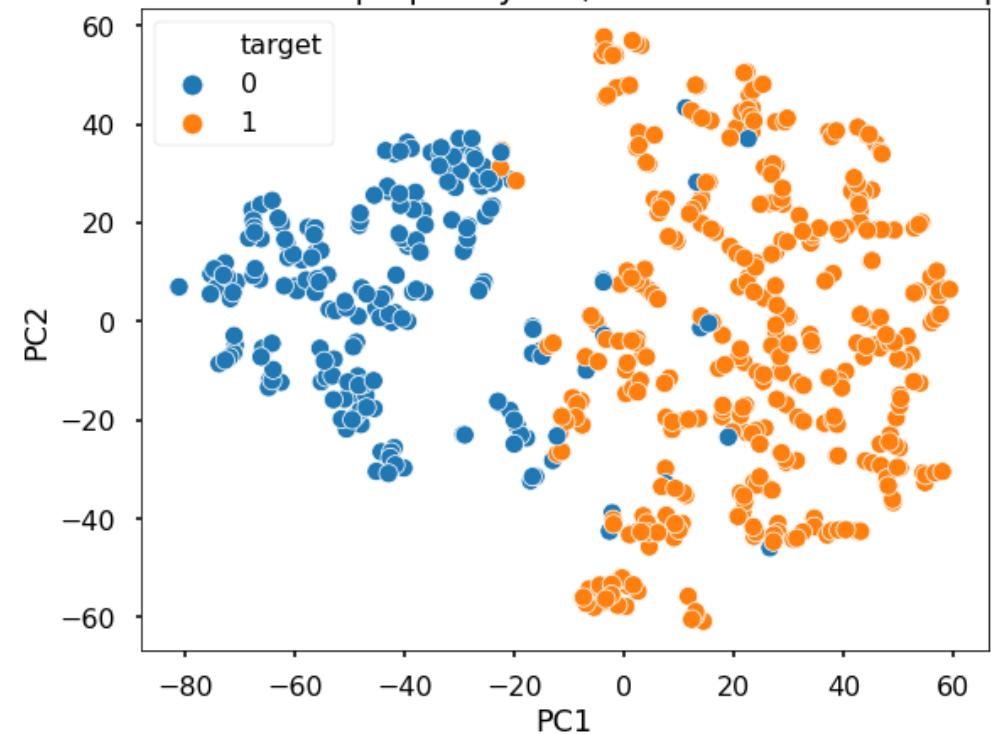
t-SNE visualization with perplexity -- 5, iterations -- 500 and epsilon -- 50



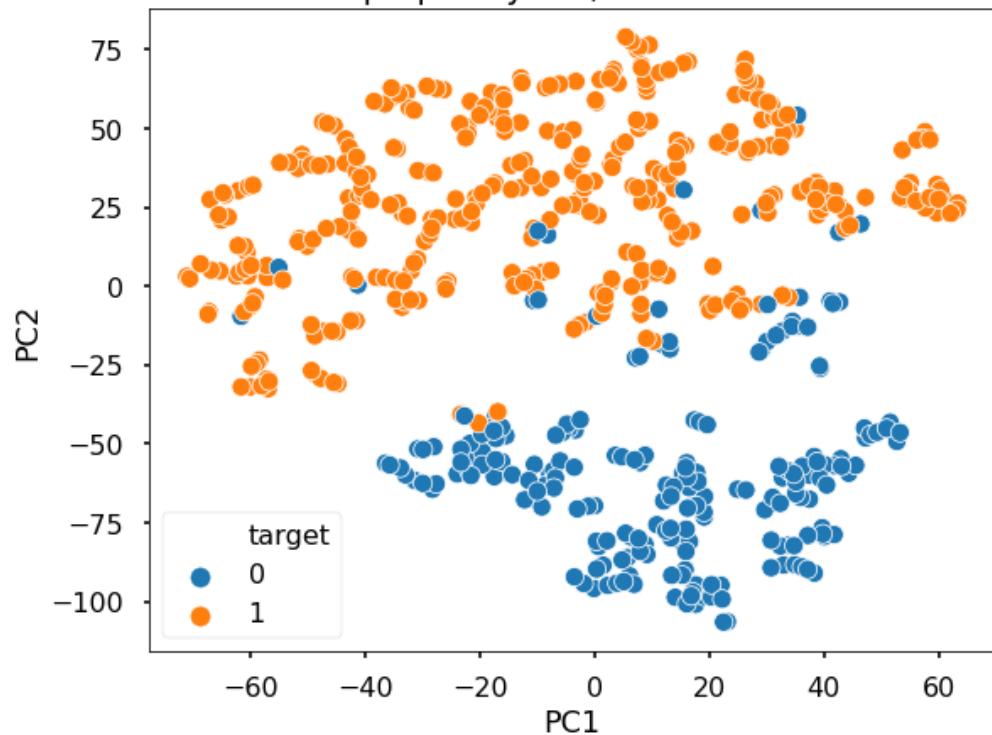
t-SNE visualization with perplexity -- 5, iterations -- 750 and epsilon -- 50



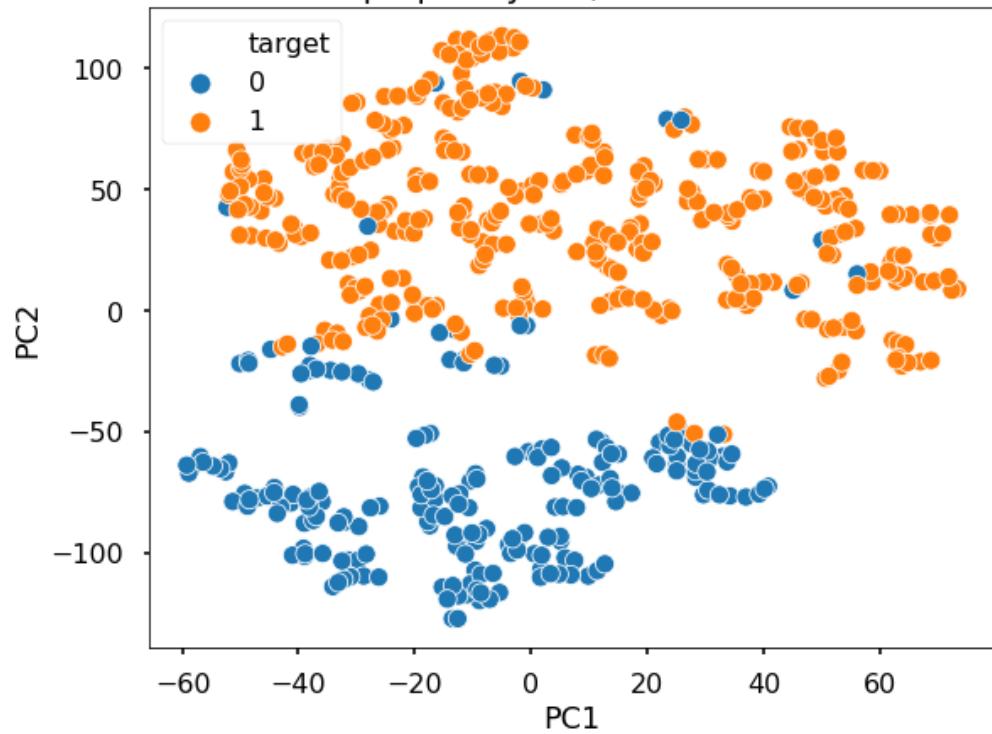
t-SNE visualization with perplexity -- 5, iterations -- 1000 and epsilon -- 50



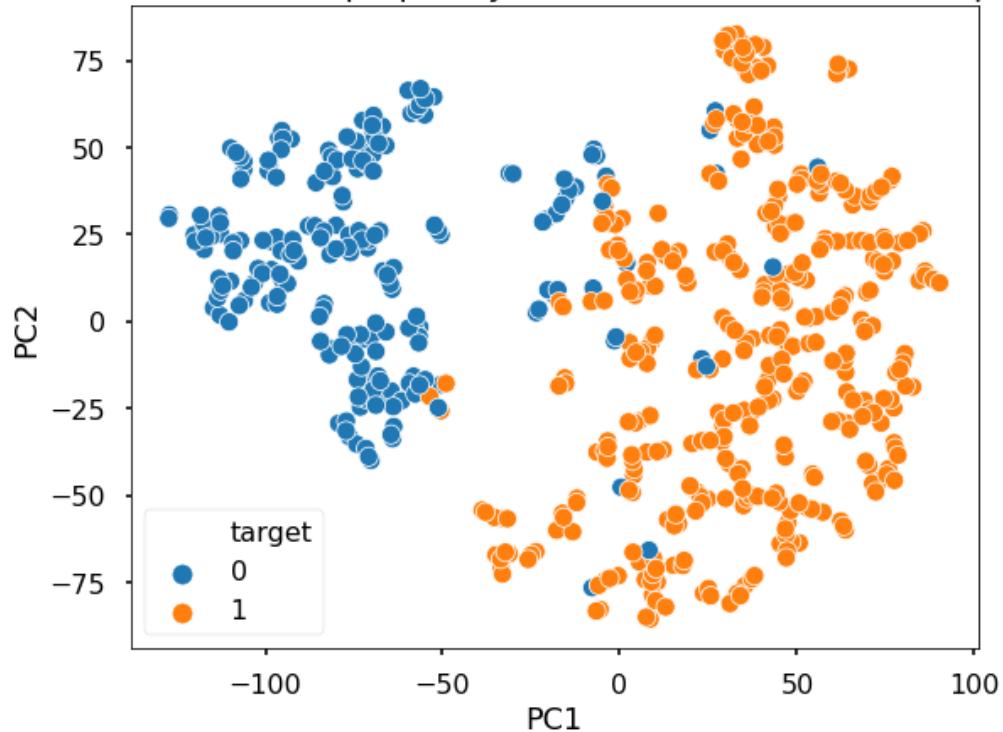
t-SNE visualization with perplexity -- 5, iterations -- 2000 and epsilon -- 50



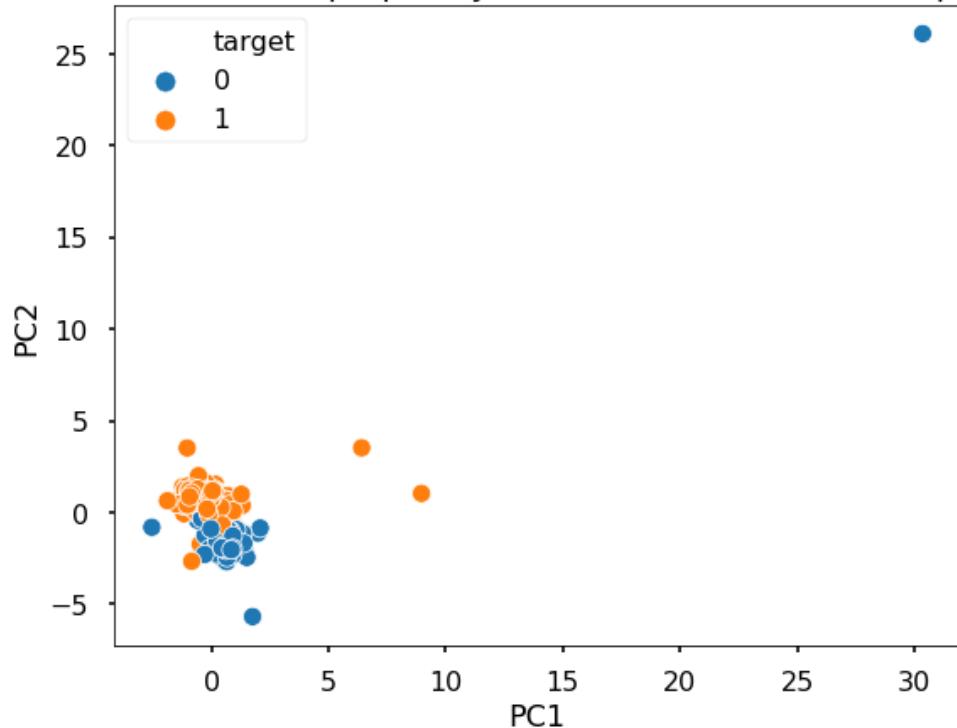
t-SNE visualization with perplexity -- 5, iterations -- 3000 and epsilon -- 50



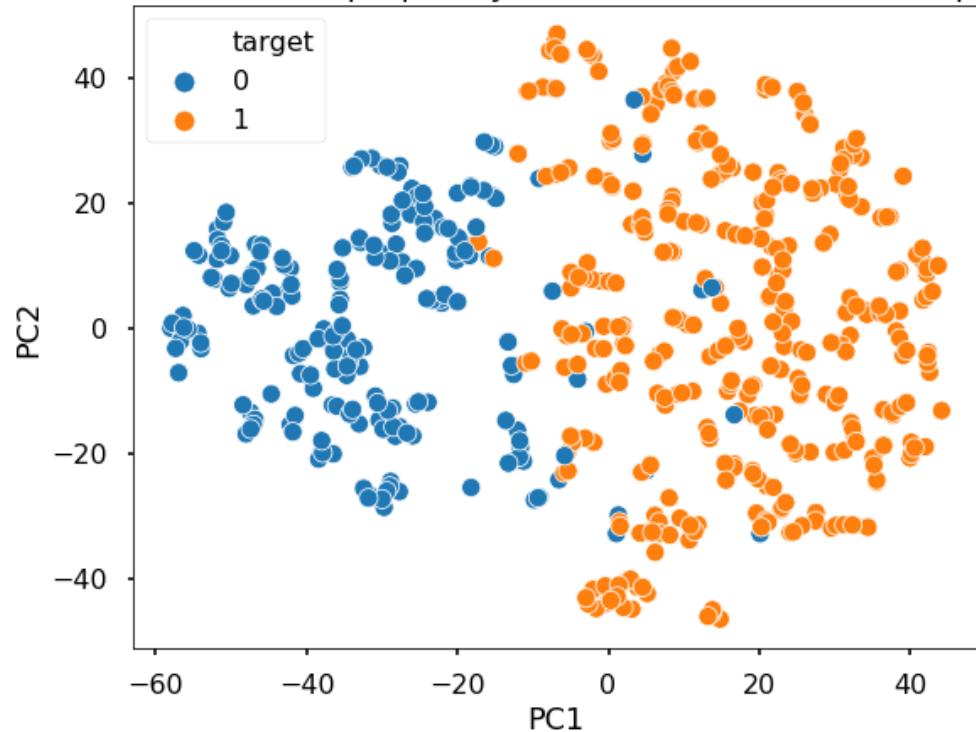
t-SNE visualization with perplexity -- 5, iterations -- 5000 and epsilon -- 50



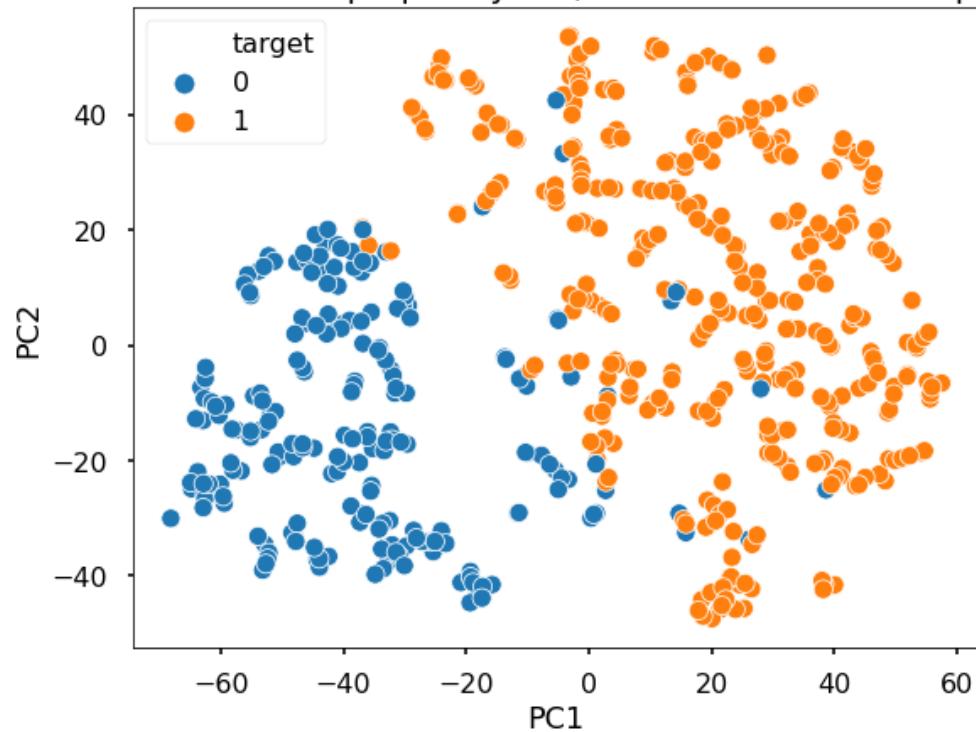
t-SNE visualization with perplexity -- 5, iterations -- 250 and epsilon -- 100



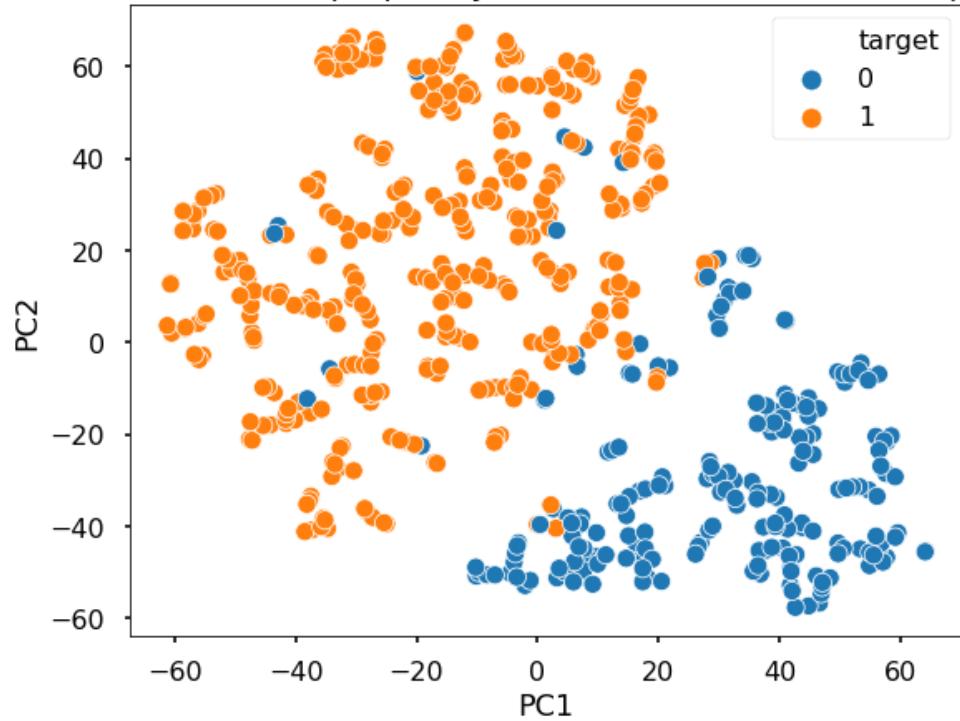
t-SNE visualization with perplexity -- 5, iterations -- 500 and epsilon -- 100



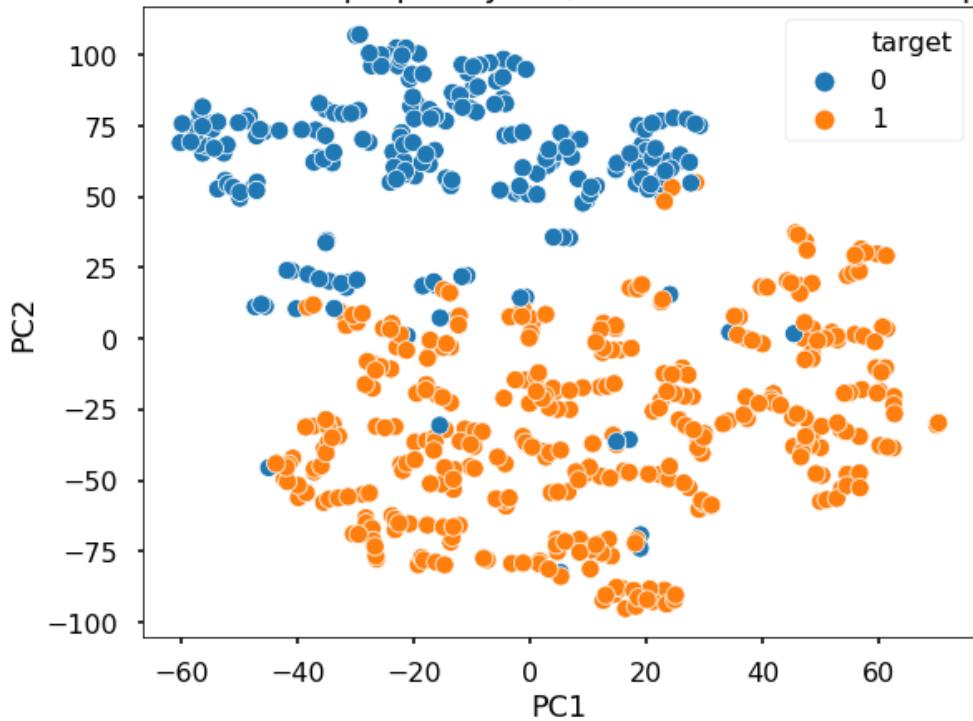
t-SNE visualization with perplexity -- 5, iterations -- 750 and epsilon -- 100



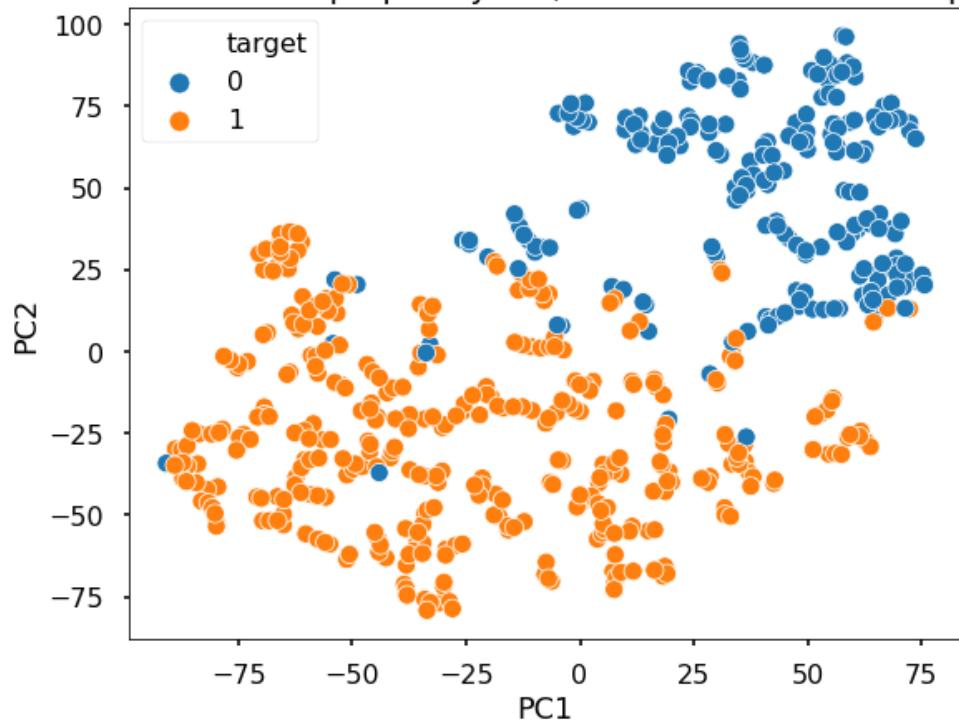
t-SNE visualization with perplexity -- 5, iterations -- 1000 and epsilon -- 100



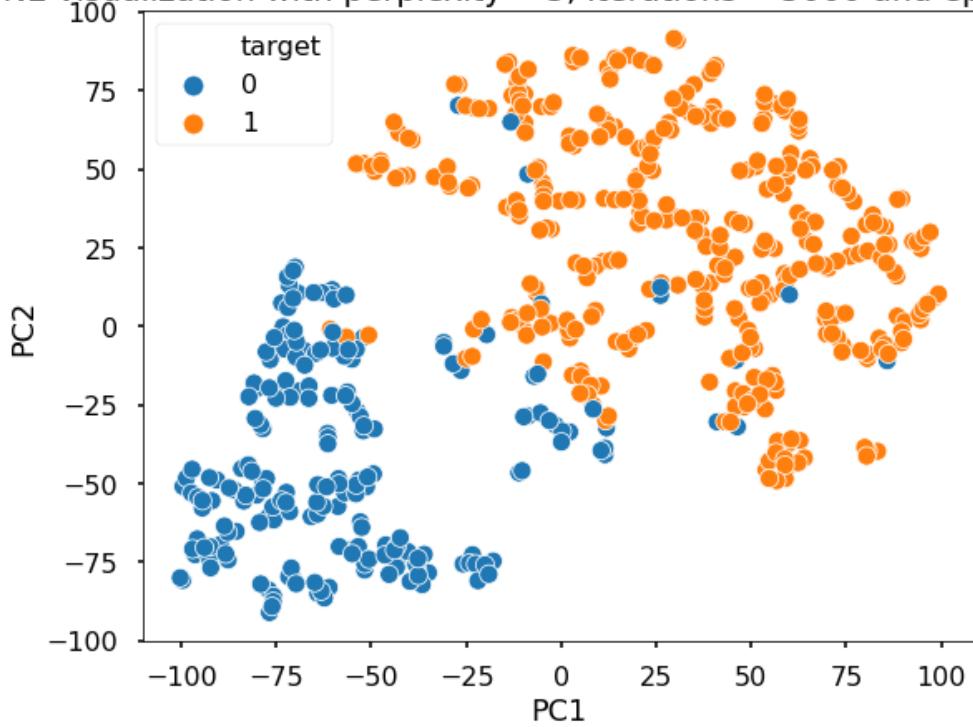
t-SNE visualization with perplexity -- 5, iterations -- 2000 and epsilon -- 100



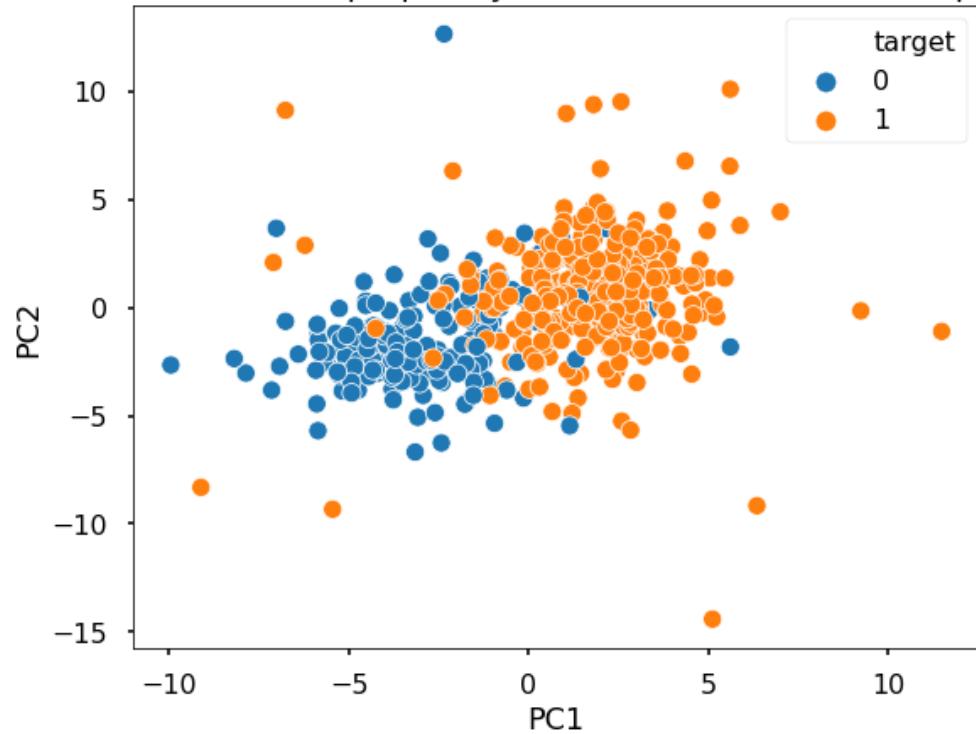
t-SNE visualization with perplexity -- 5, iterations -- 3000 and epsilon -- 100



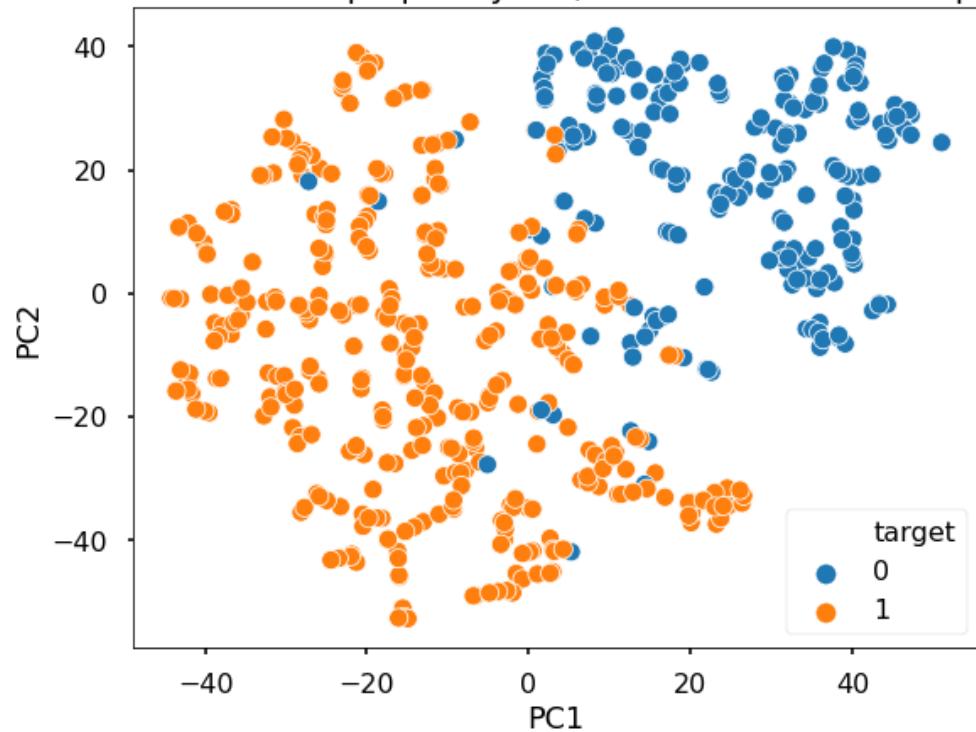
t-SNE visualization with perplexity -- 5, iterations -- 5000 and epsilon -- 100



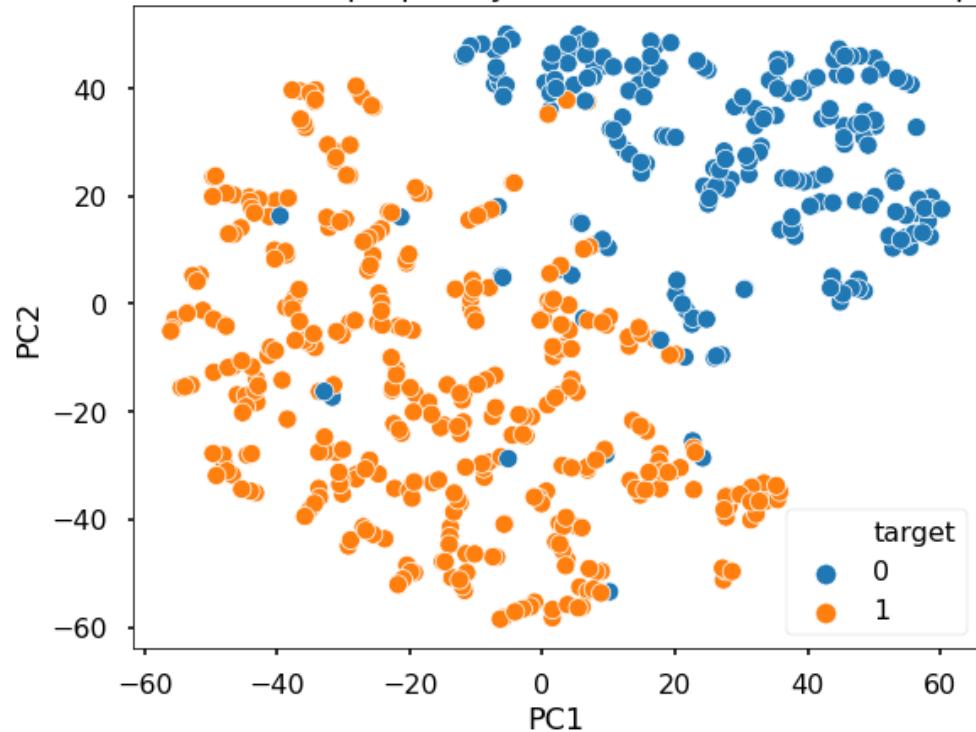
t-SNE visualization with perplexity -- 5, iterations -- 250 and epsilon -- 200



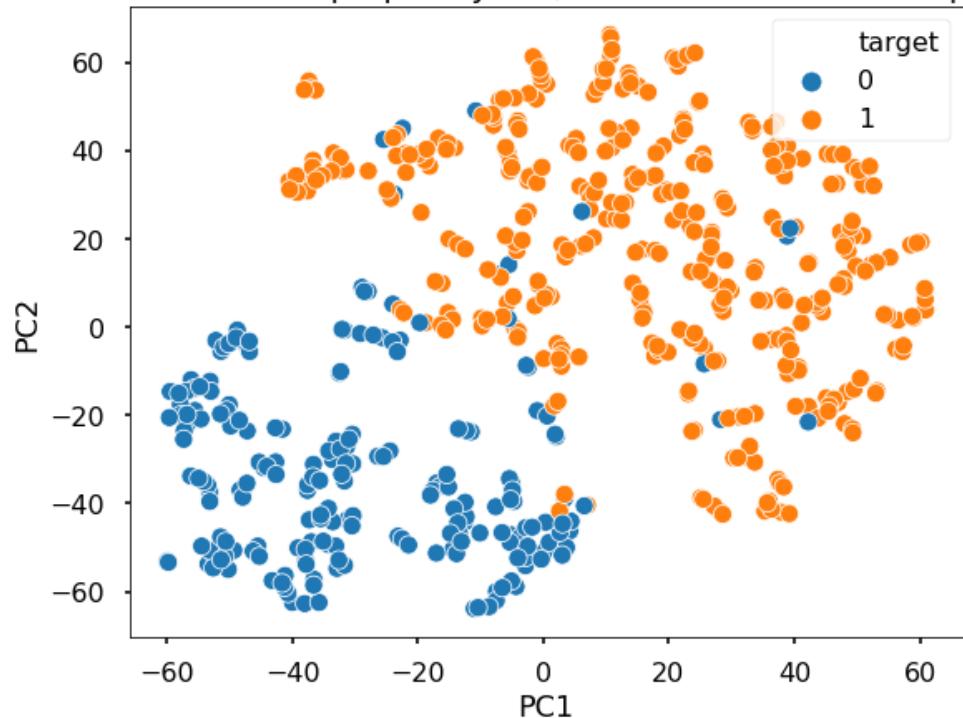
t-SNE visualization with perplexity -- 5, iterations -- 500 and epsilon -- 200



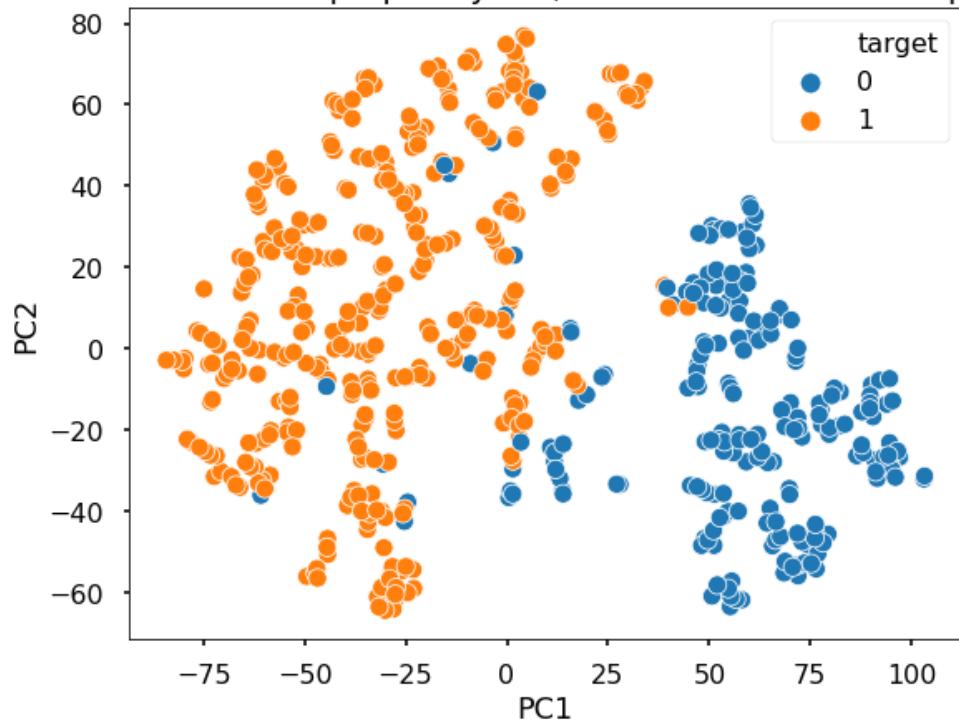
t-SNE visualization with perplexity -- 5, iterations -- 750 and epsilon -- 200



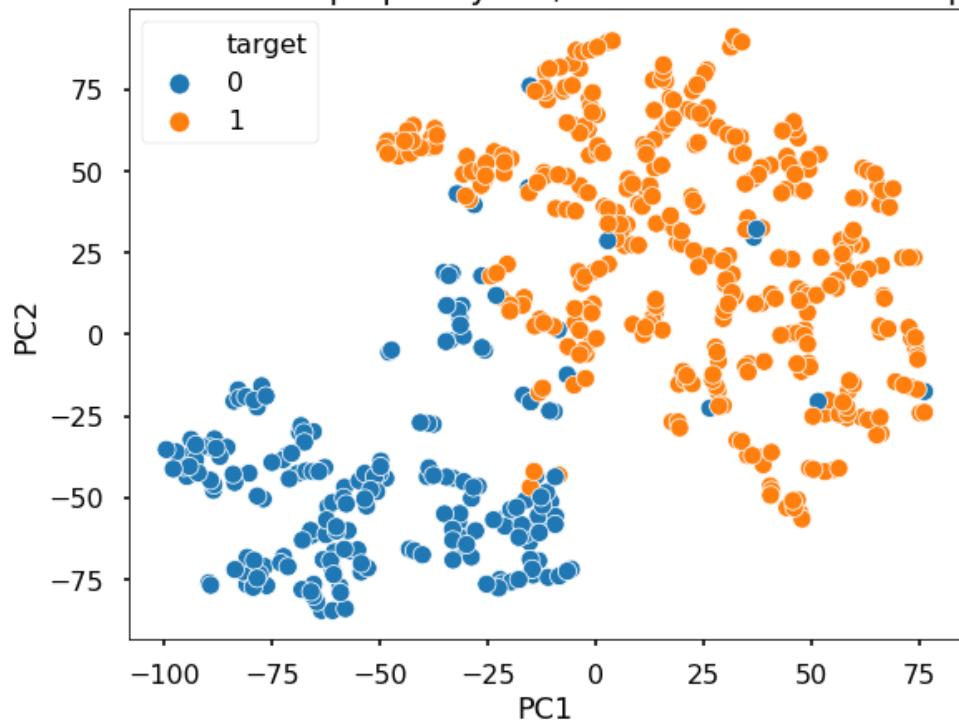
t-SNE visualization with perplexity -- 5, iterations -- 1000 and epsilon -- 200



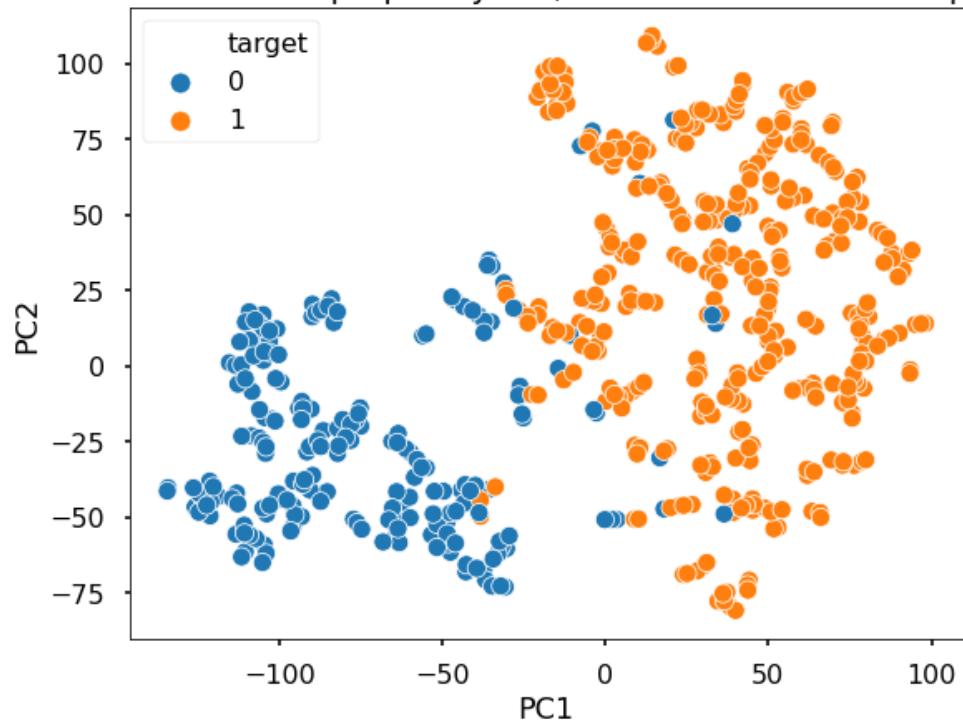
t-SNE visualization with perplexity -- 5, iterations -- 2000 and epsilon -- 200



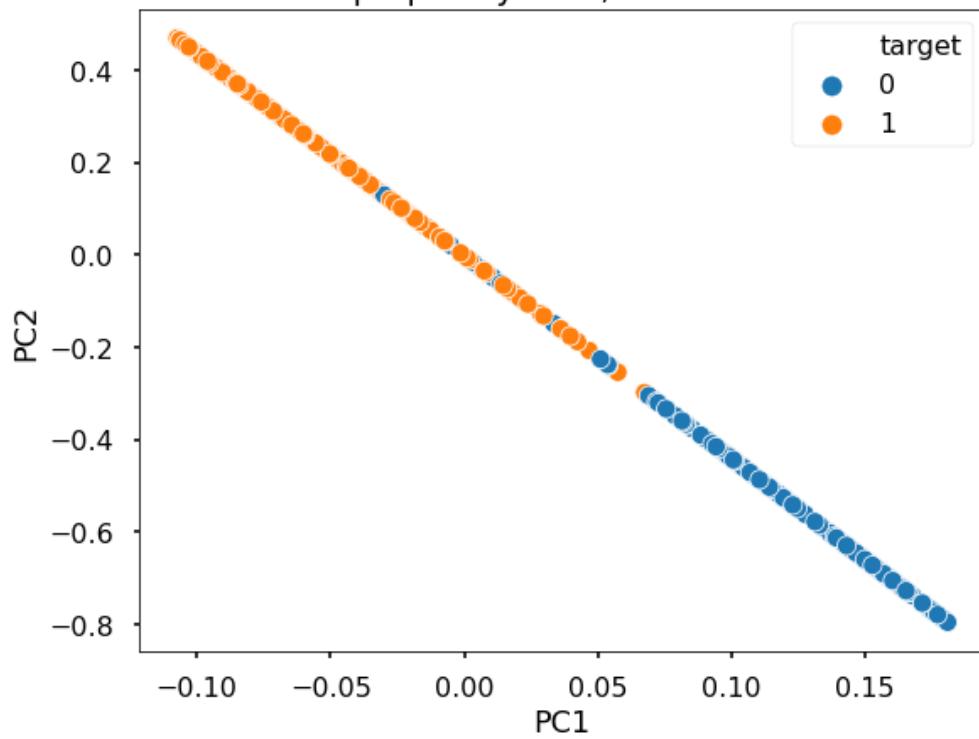
t-SNE visualization with perplexity -- 5, iterations -- 3000 and epsilon -- 200



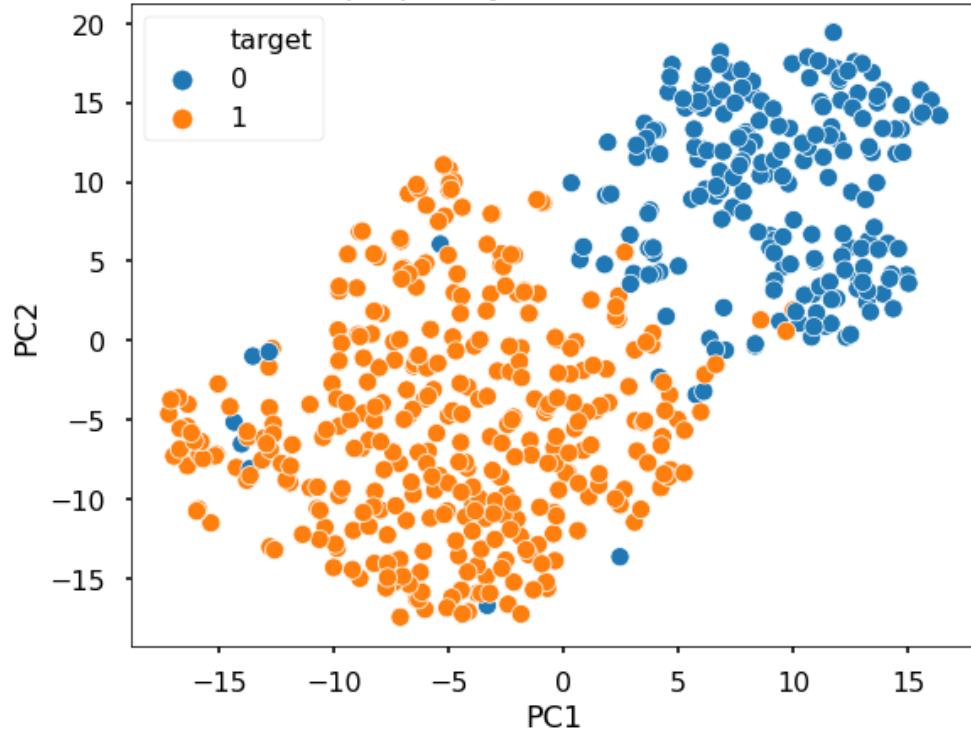
t-SNE visualization with perplexity -- 5, iterations -- 5000 and epsilon -- 200



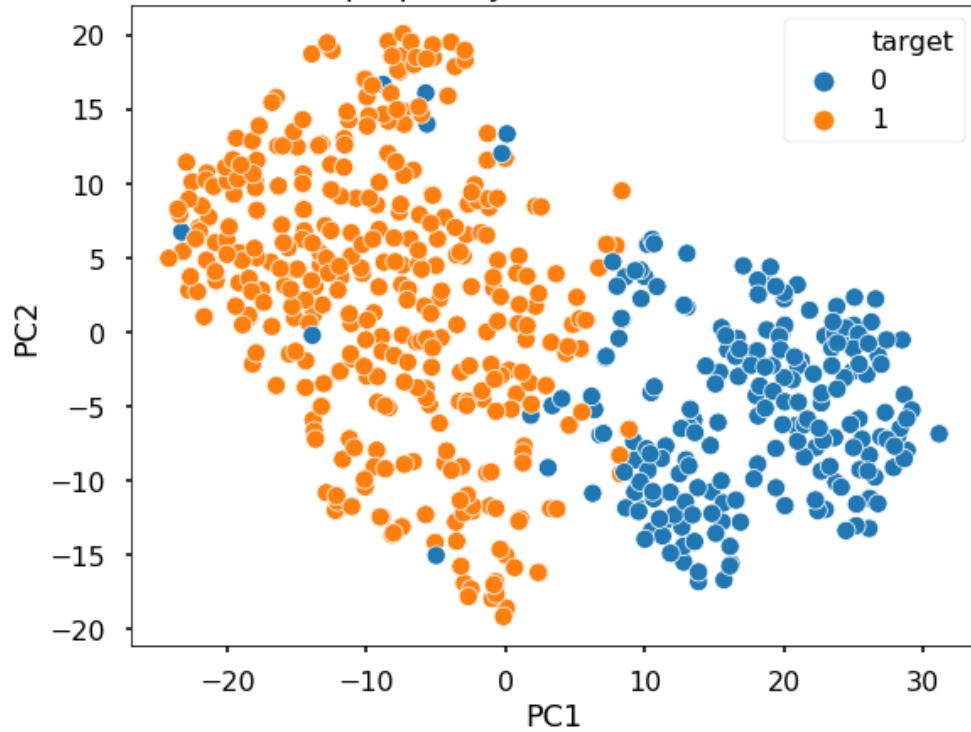
t-SNE visualization with perplexity -- 30, iterations -- 250 and epsilon -- 10



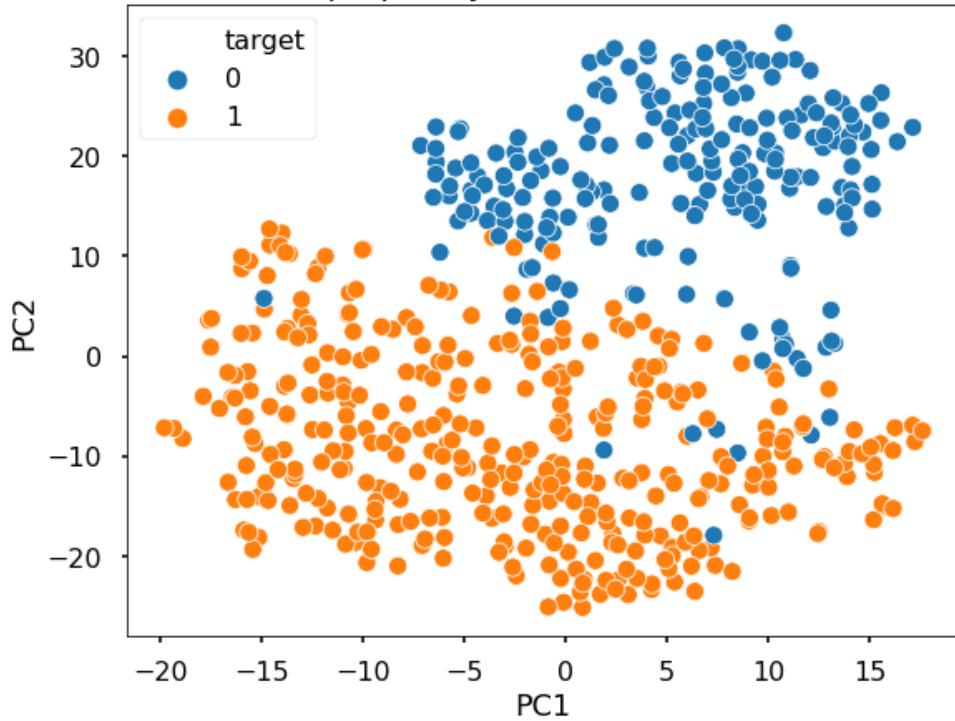
t-SNE visualization with perplexity -- 30, iterations -- 500 and epsilon -- 10



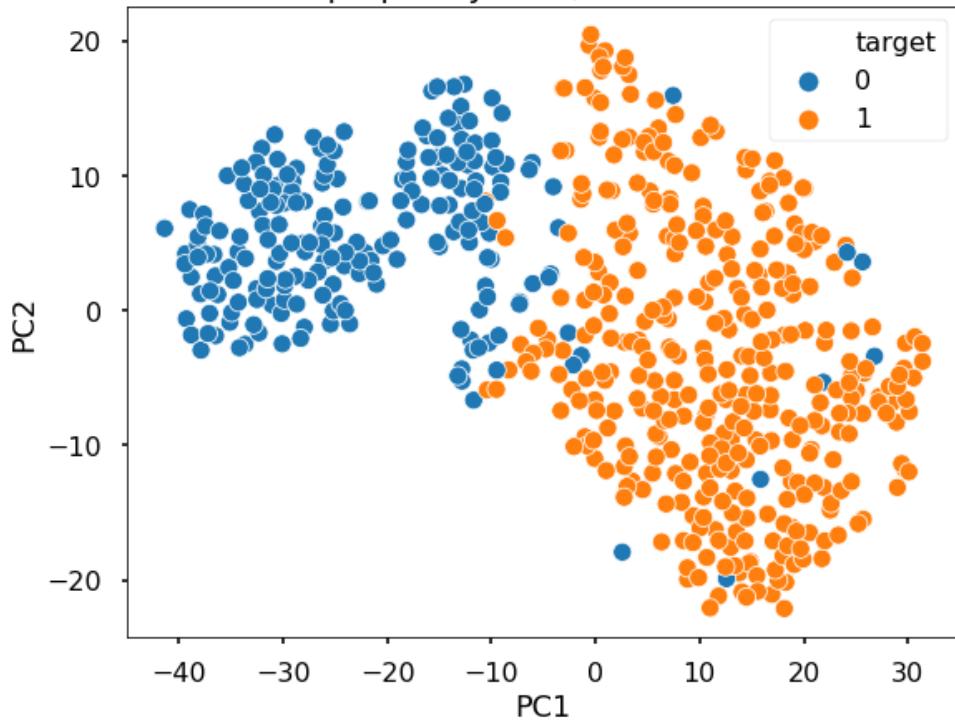
t-SNE visualization with perplexity -- 30, iterations -- 750 and epsilon -- 10



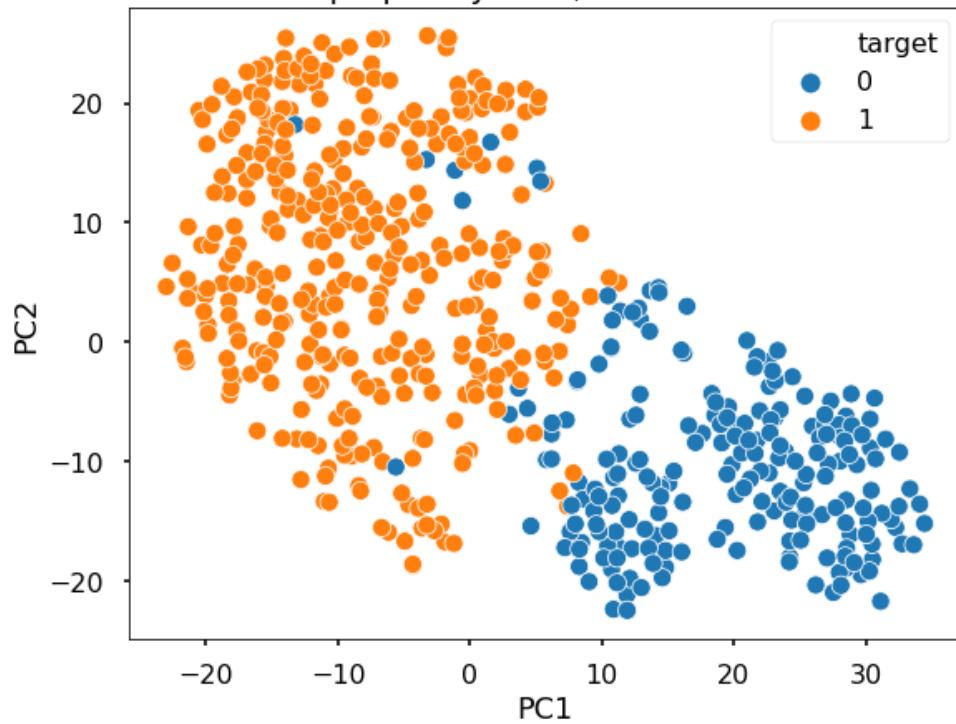
t-SNE visualization with perplexity -- 30, iterations -- 1000 and epsilon -- 10



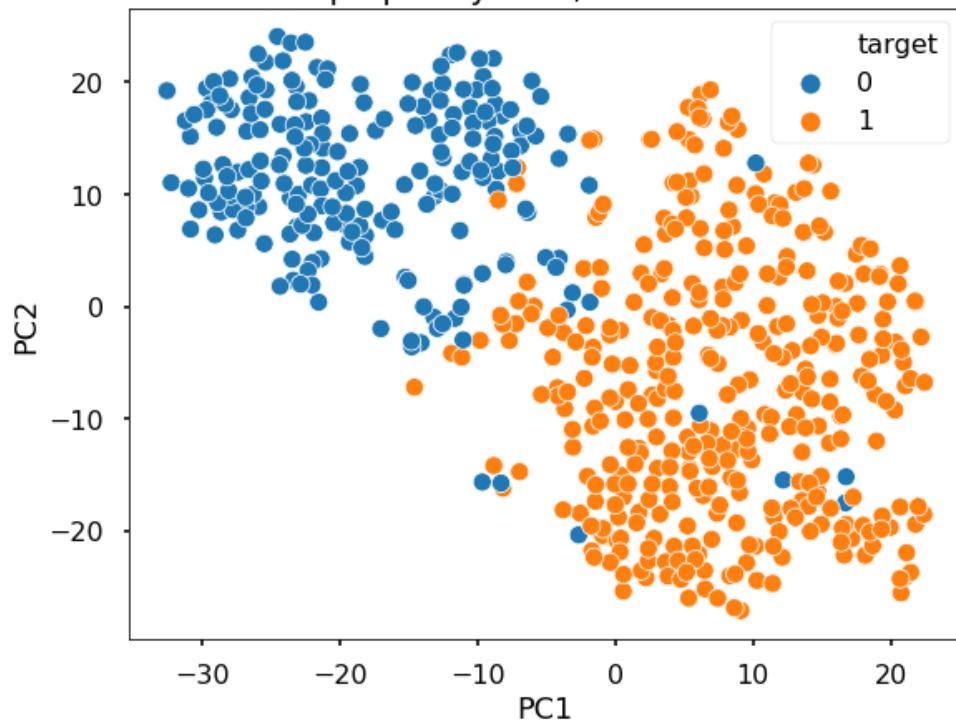
t-SNE visualization with perplexity -- 30, iterations -- 2000 and epsilon -- 10



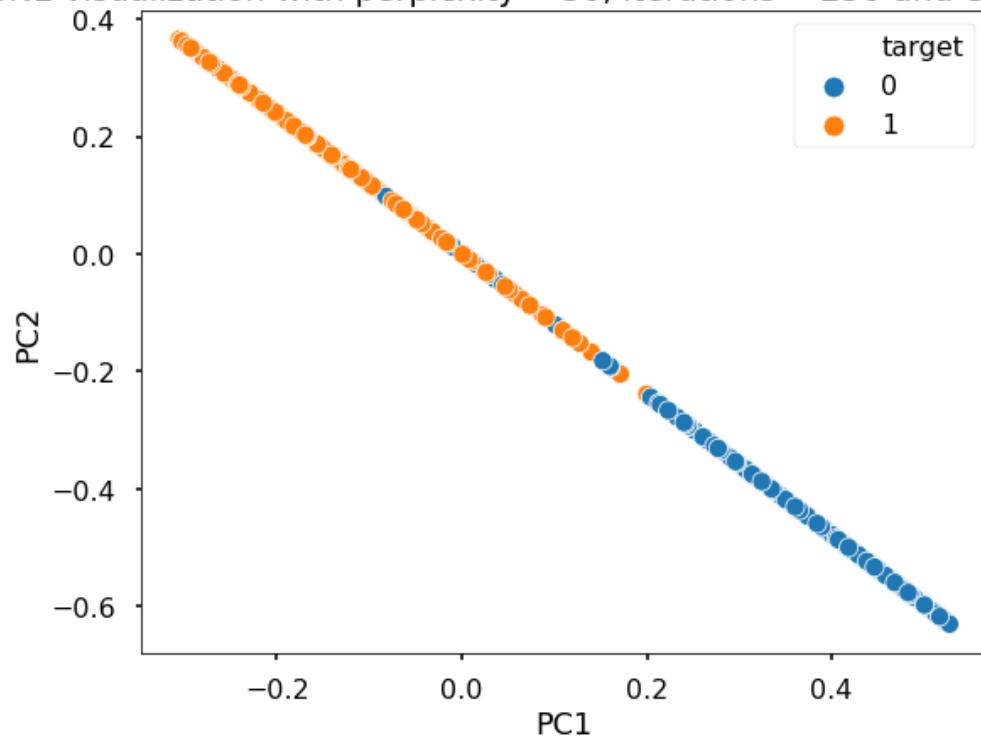
t-SNE visualization with perplexity -- 30, iterations -- 3000 and epsilon -- 10



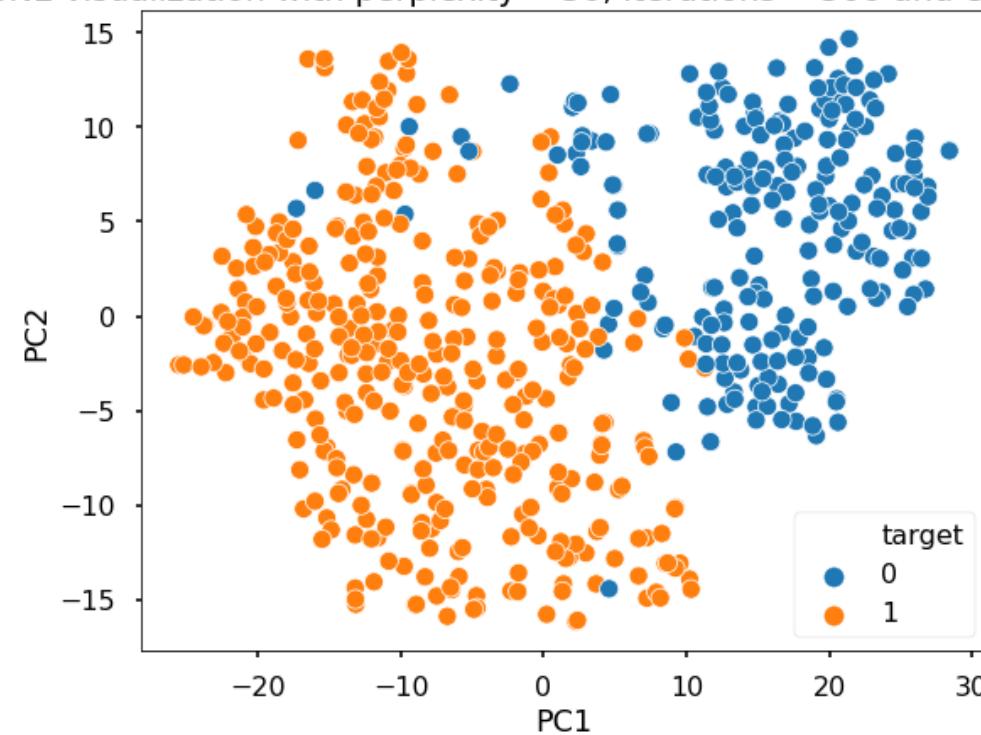
t-SNE visualization with perplexity -- 30, iterations -- 5000 and epsilon -- 10



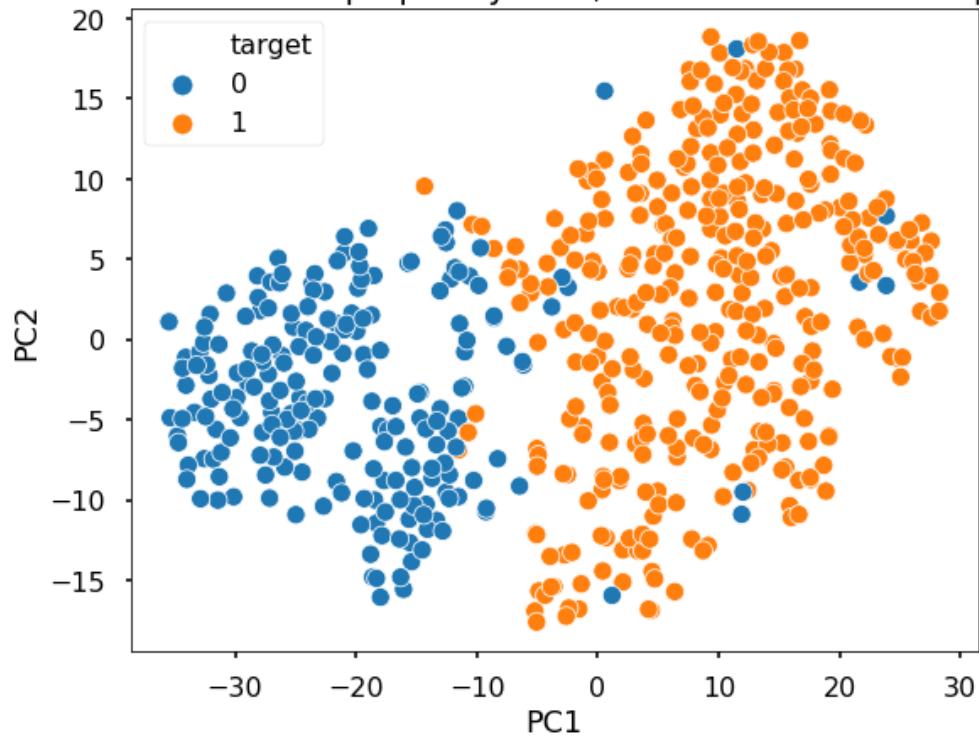
t-SNE visualization with perplexity -- 30, iterations -- 250 and epsilon -- 30



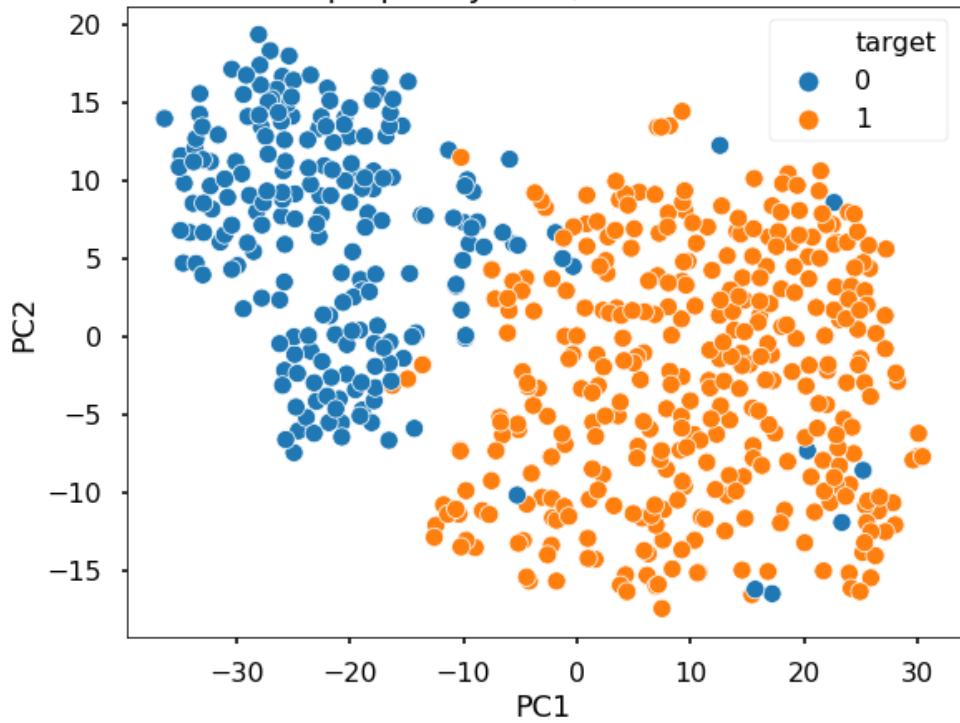
t-SNE visualization with perplexity -- 30, iterations -- 500 and epsilon -- 30



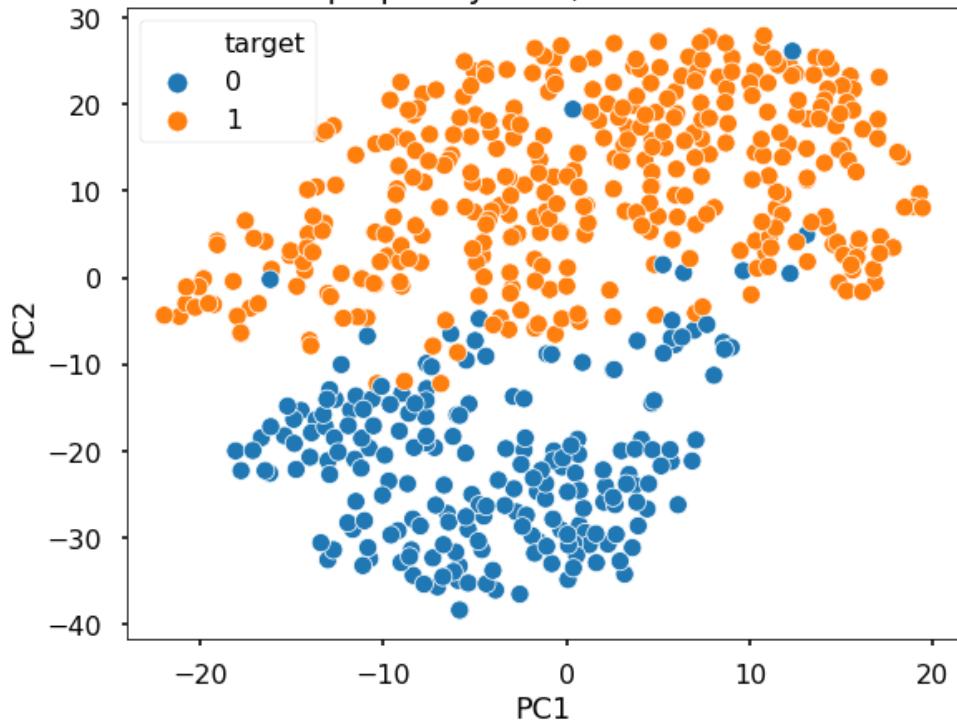
t-SNE visualization with perplexity -- 30, iterations -- 750 and epsilon -- 30



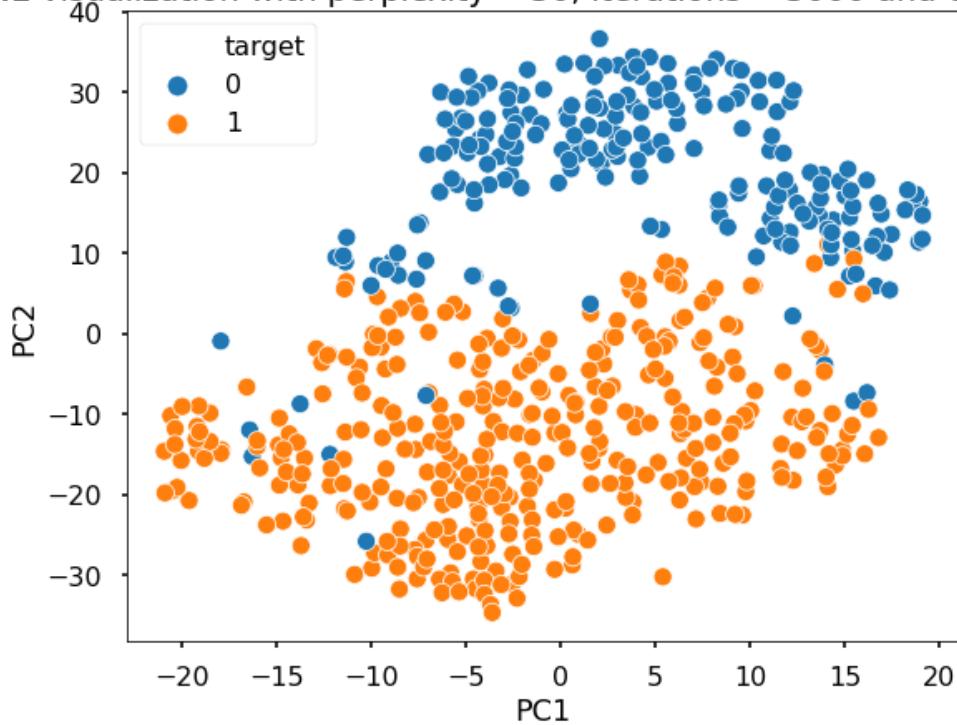
t-SNE visualization with perplexity -- 30, iterations -- 1000 and epsilon -- 30



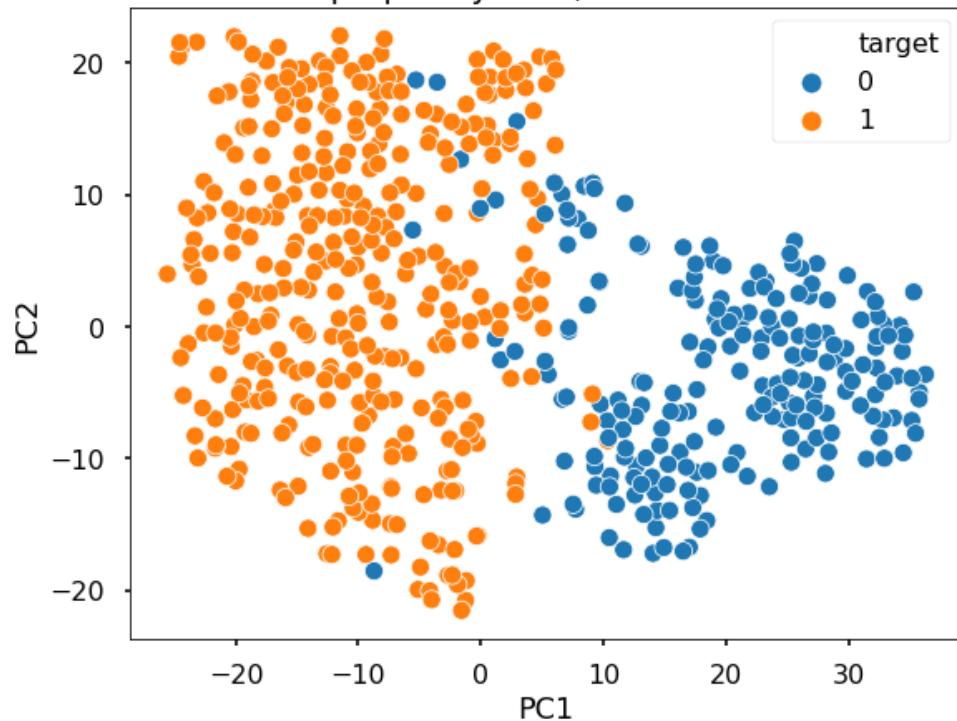
t-SNE visualization with perplexity -- 30, iterations -- 2000 and epsilon -- 30



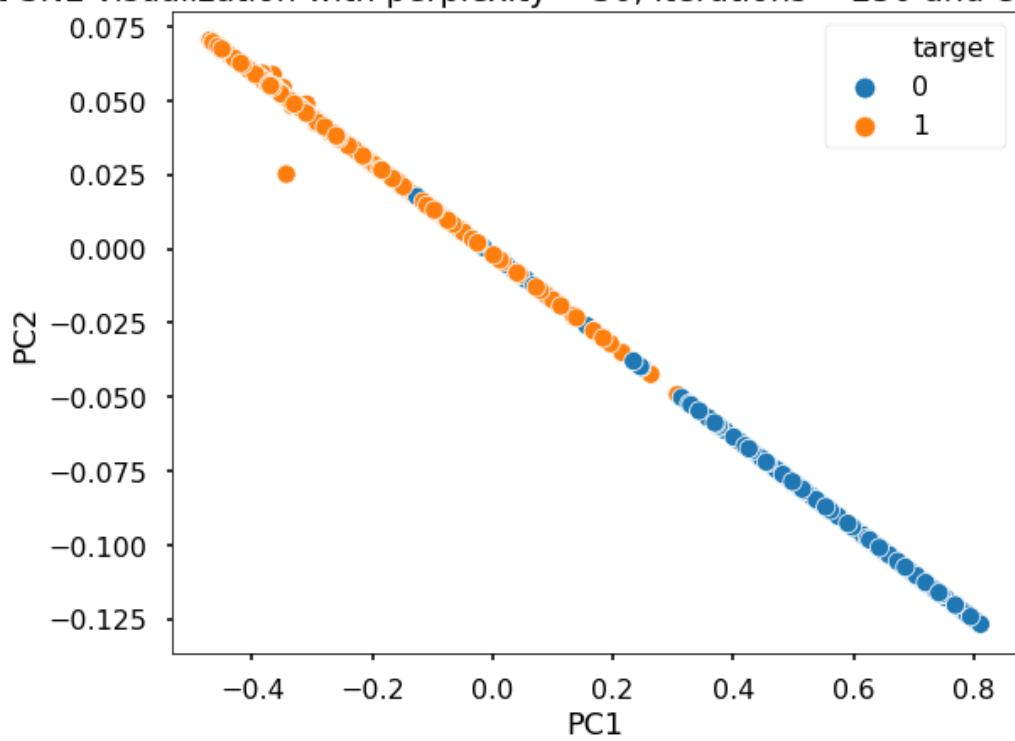
t-SNE visualization with perplexity -- 30, iterations -- 3000 and epsilon -- 30



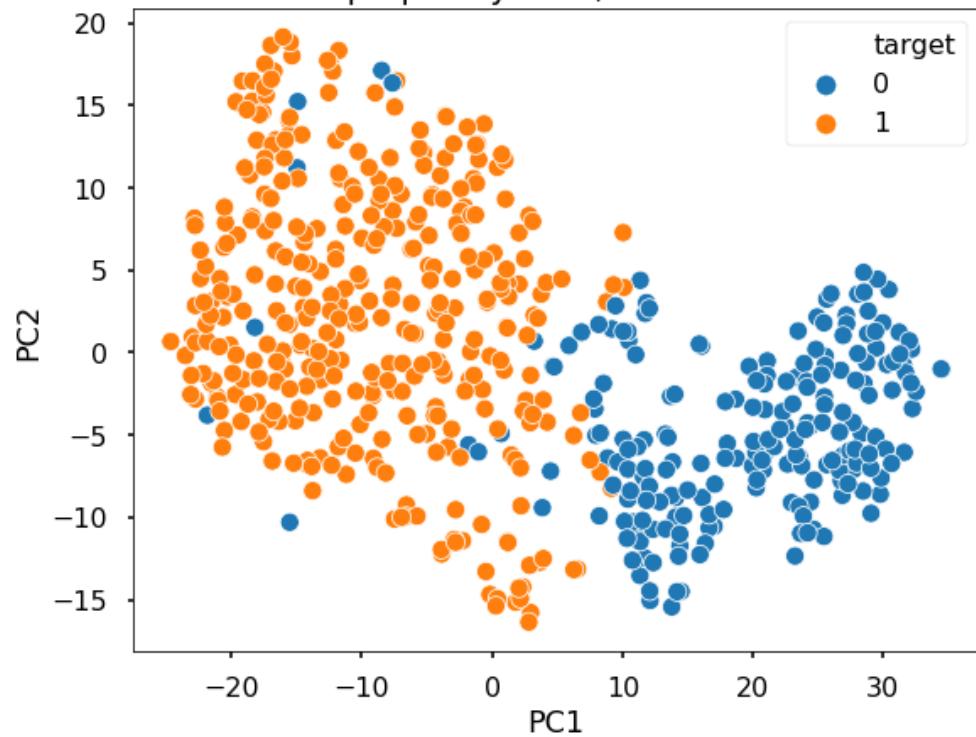
t-SNE visualization with perplexity -- 30, iterations -- 5000 and epsilon -- 30



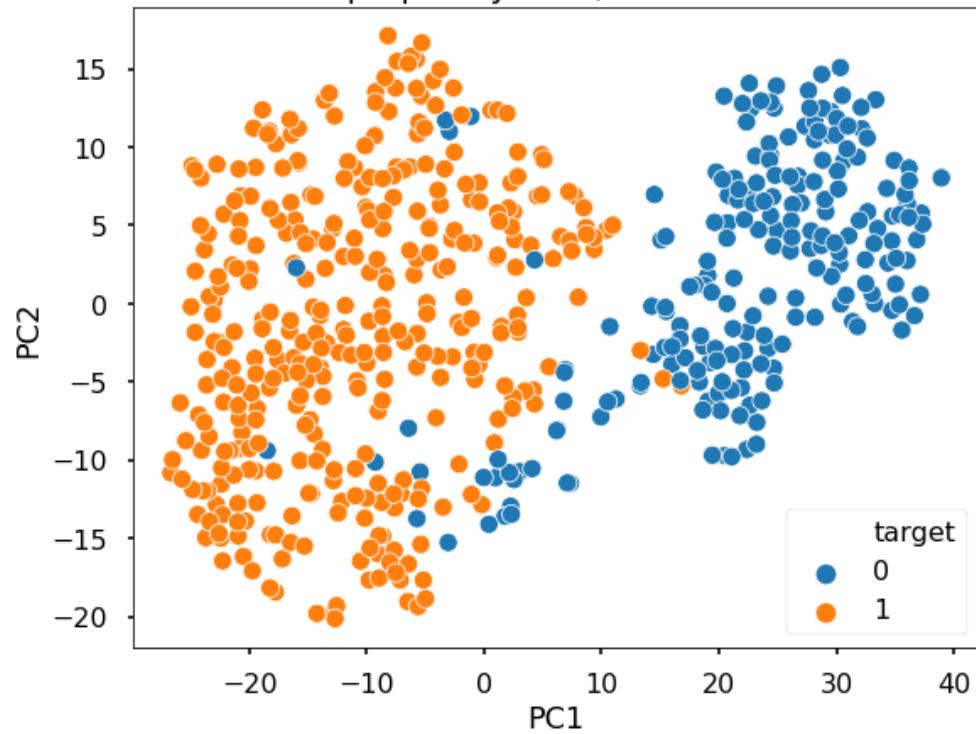
t-SNE visualization with perplexity -- 30, iterations -- 250 and epsilon -- 50



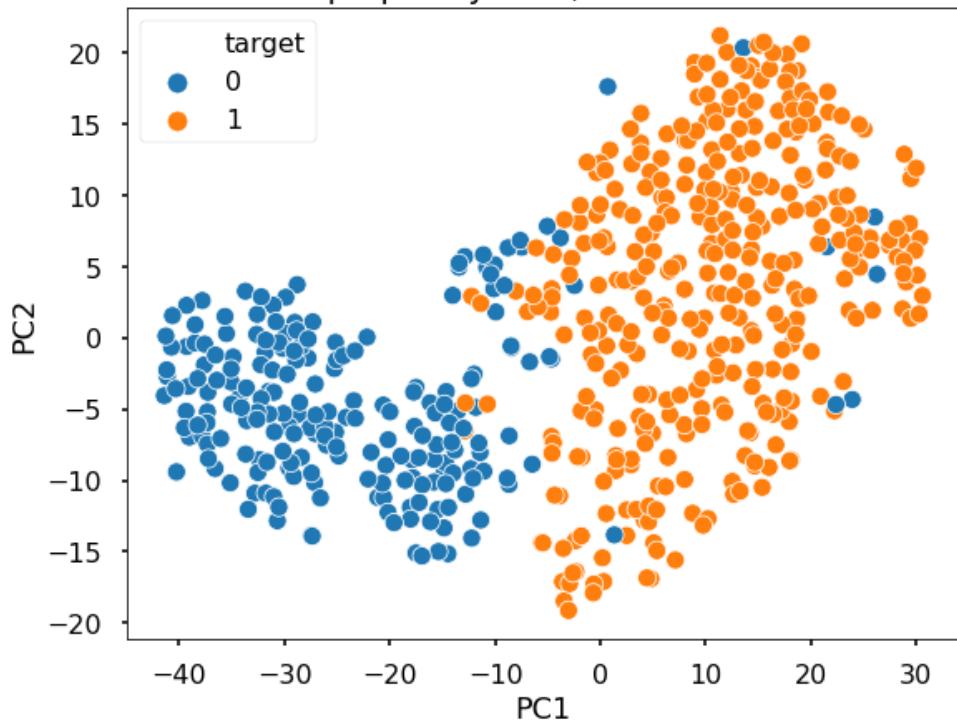
t-SNE visualization with perplexity -- 30, iterations -- 500 and epsilon -- 50



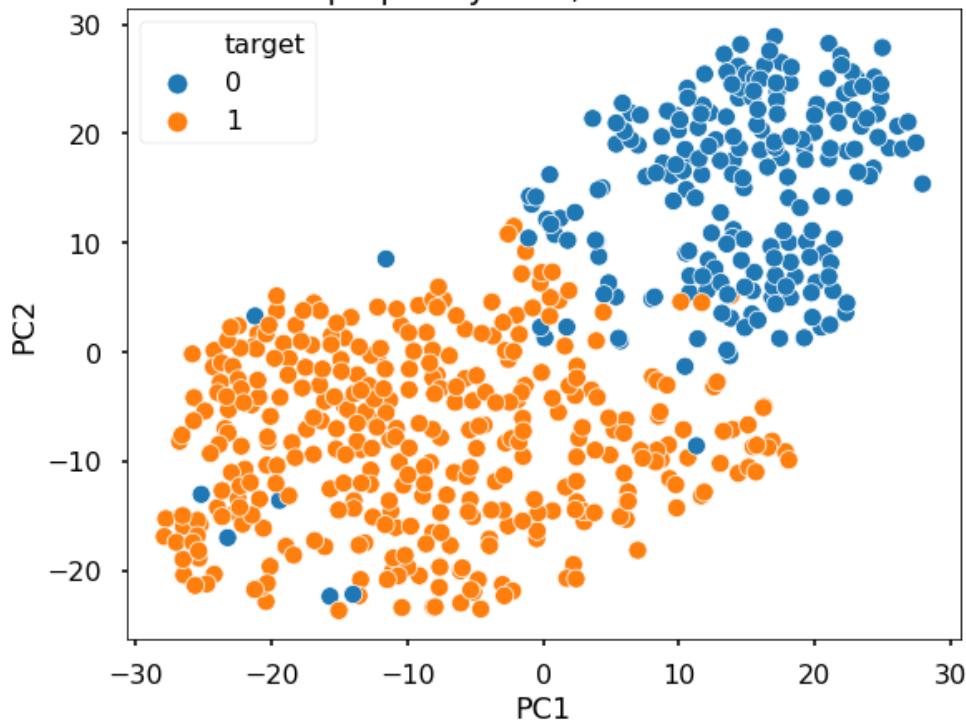
t-SNE visualization with perplexity -- 30, iterations -- 750 and epsilon -- 50



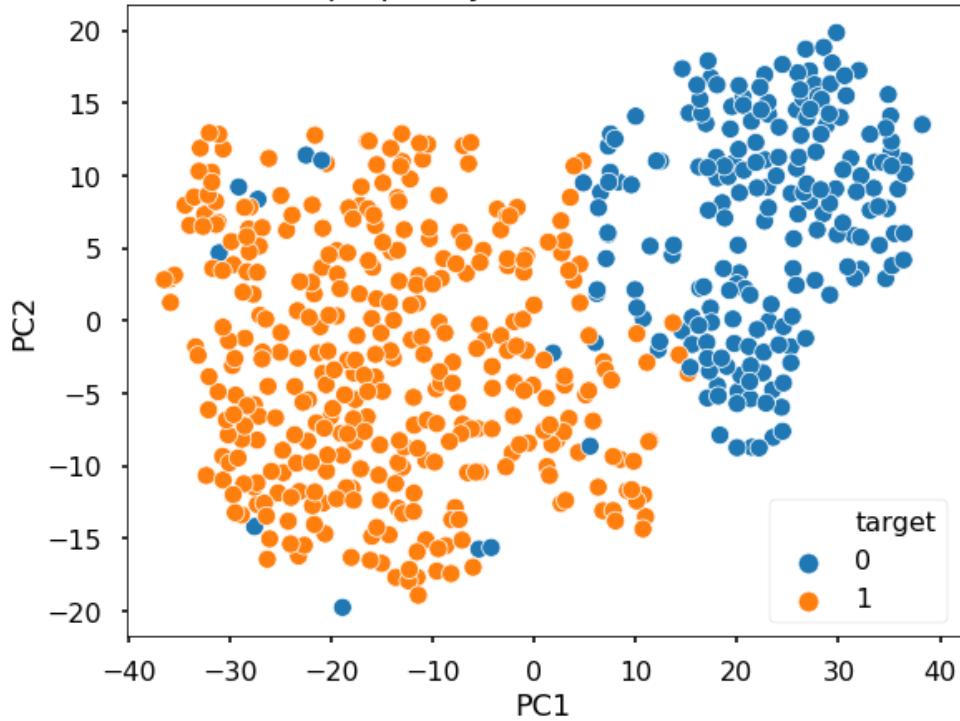
t-SNE visualization with perplexity -- 30, iterations -- 1000 and epsilon -- 50



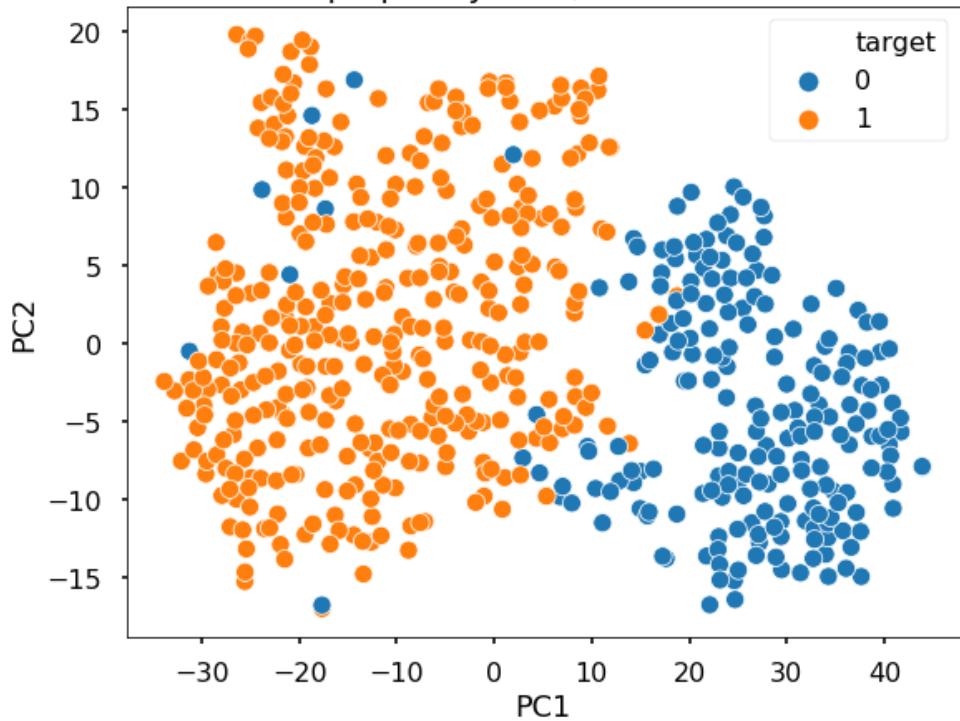
t-SNE visualization with perplexity -- 30, iterations -- 2000 and epsilon -- 50



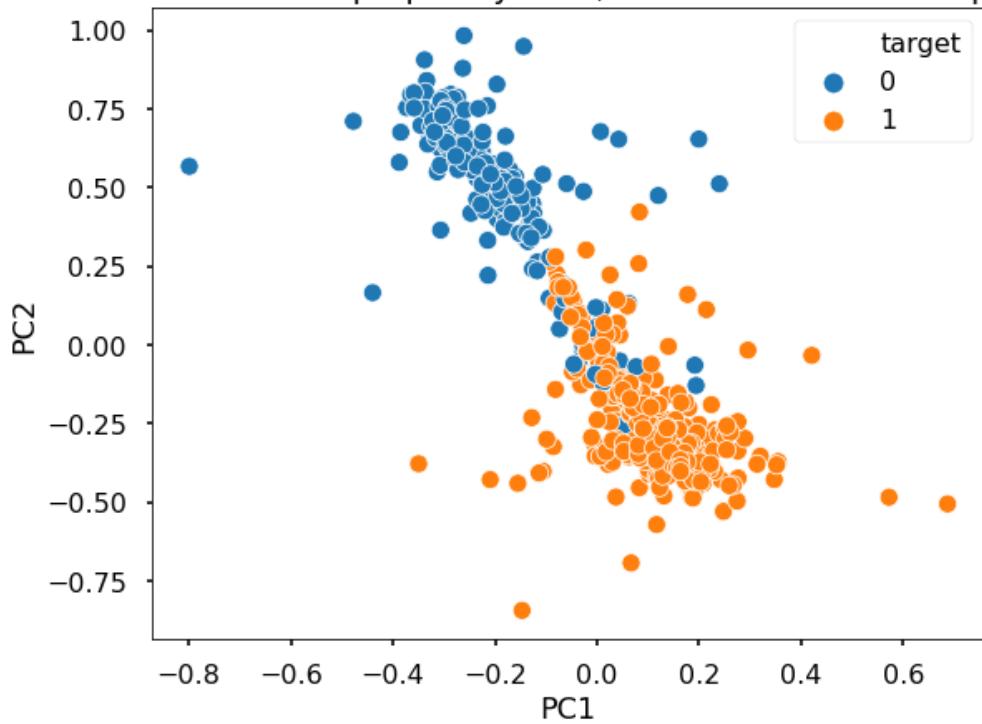
t-SNE visualization with perplexity -- 30, iterations -- 3000 and epsilon -- 50



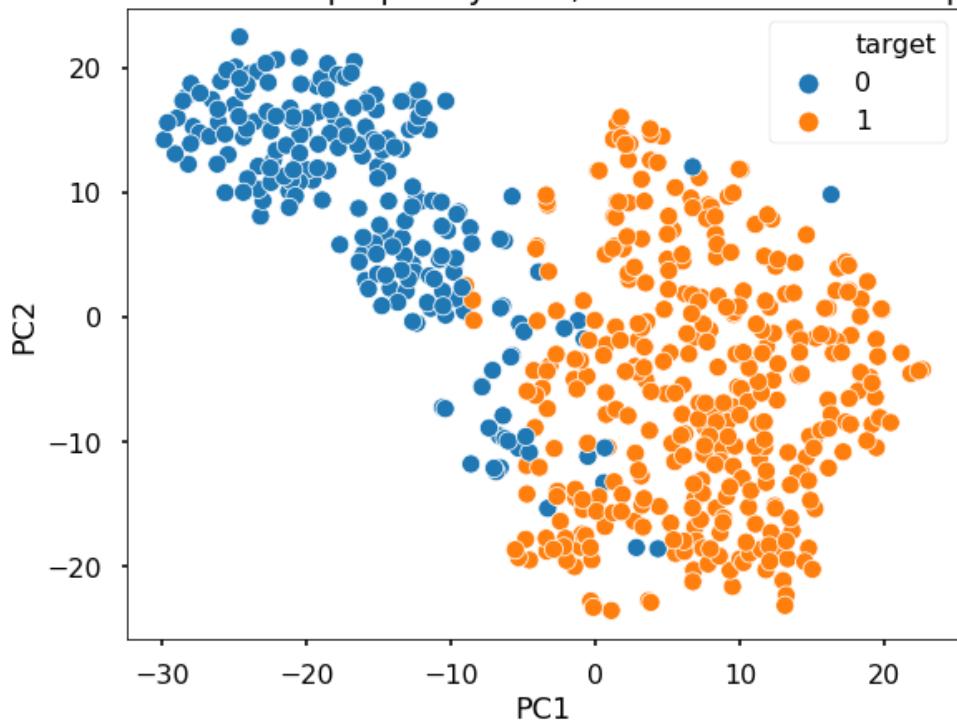
t-SNE visualization with perplexity -- 30, iterations -- 5000 and epsilon -- 50



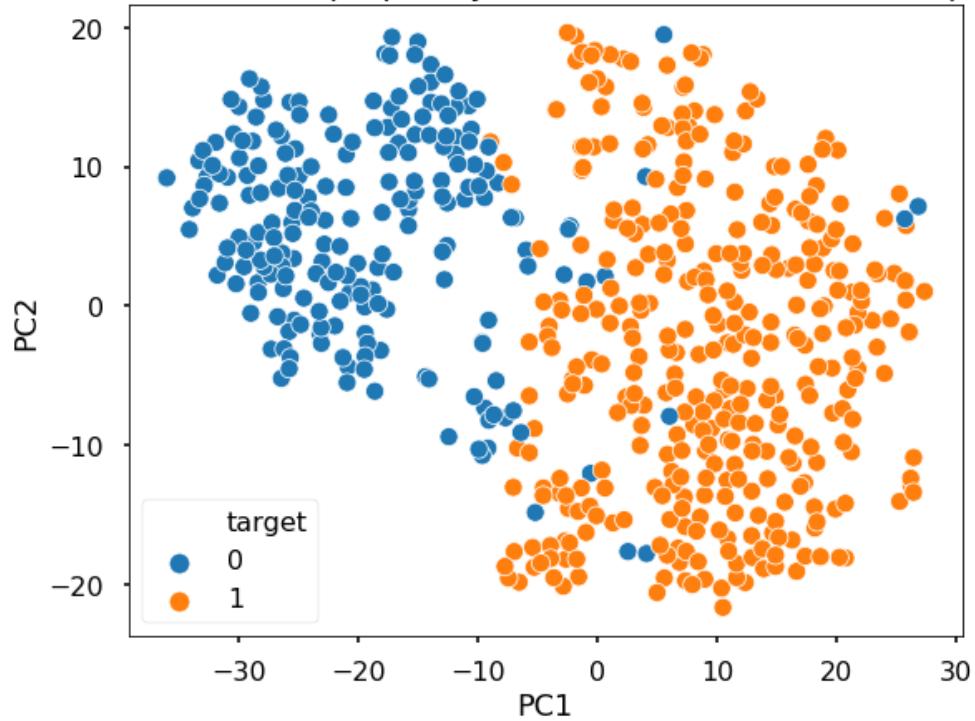
t-SNE visualization with perplexity -- 30, iterations -- 250 and epsilon -- 100



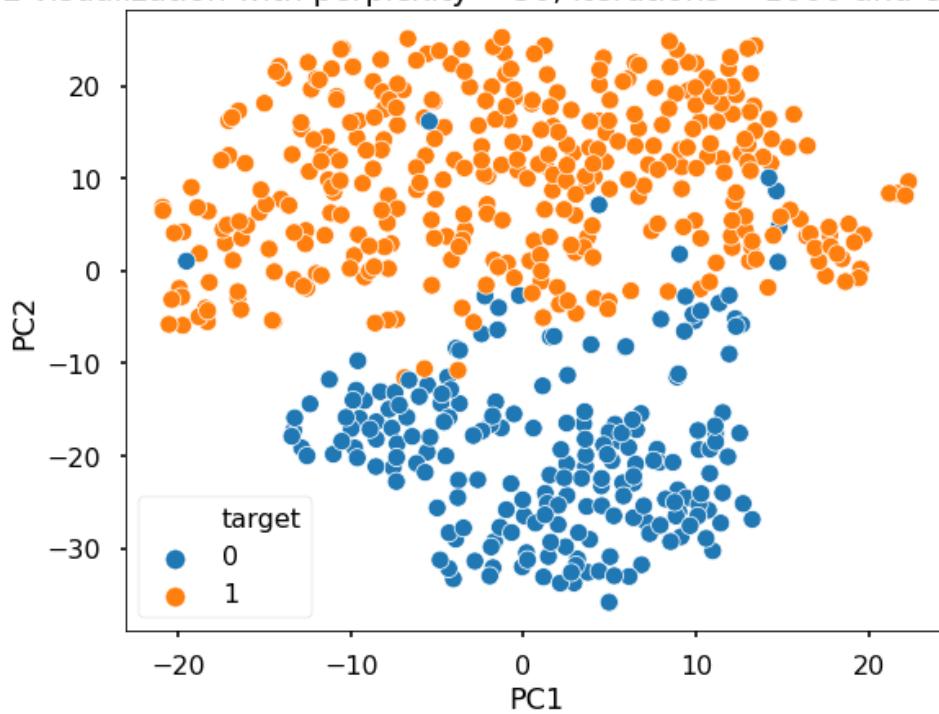
t-SNE visualization with perplexity -- 30, iterations -- 500 and epsilon -- 100



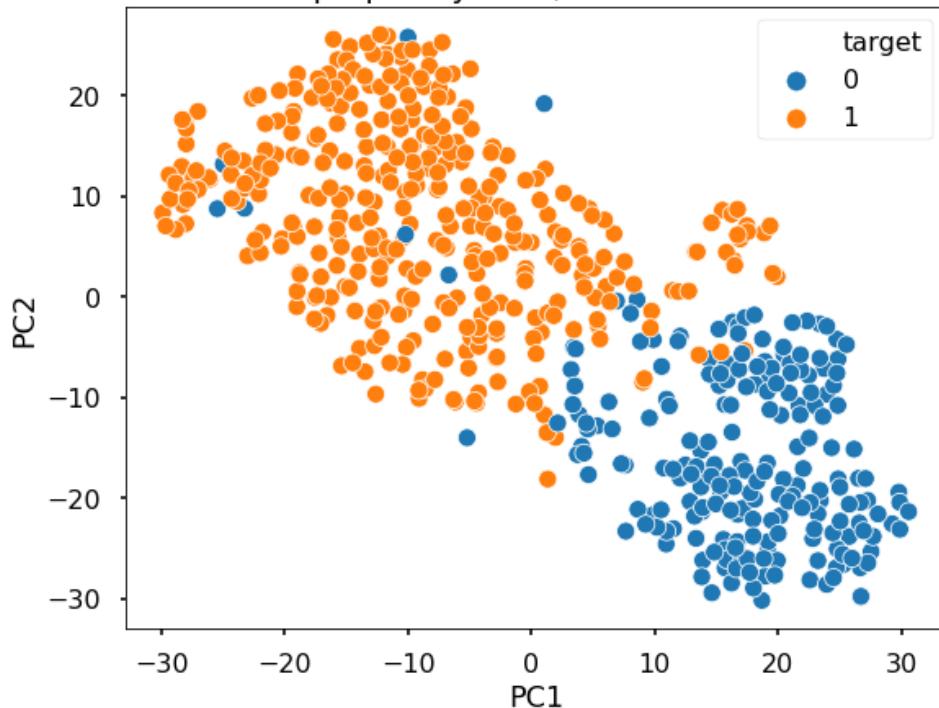
t-SNE visualization with perplexity -- 30, iterations -- 750 and epsilon -- 100



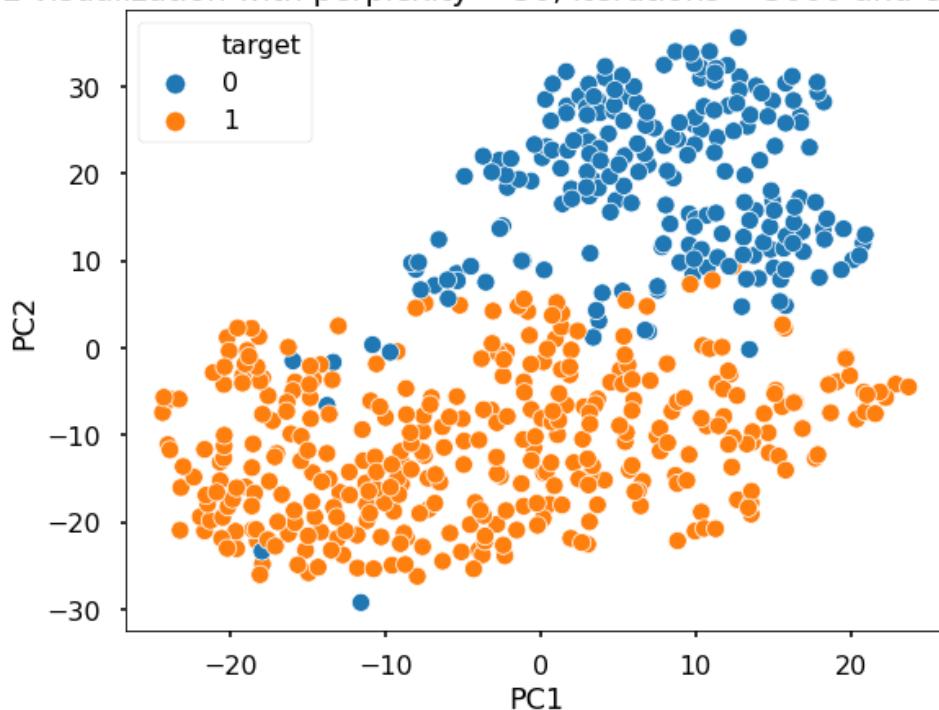
t-SNE visualization with perplexity -- 30, iterations -- 1000 and epsilon -- 100



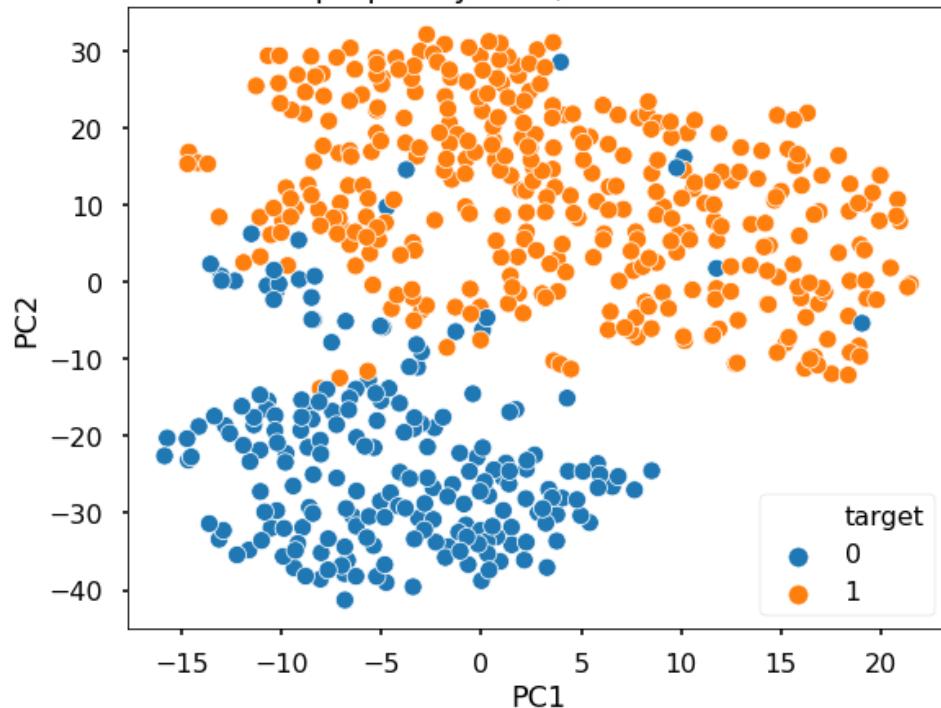
t-SNE visualization with perplexity -- 30, iterations -- 2000 and epsilon -- 100



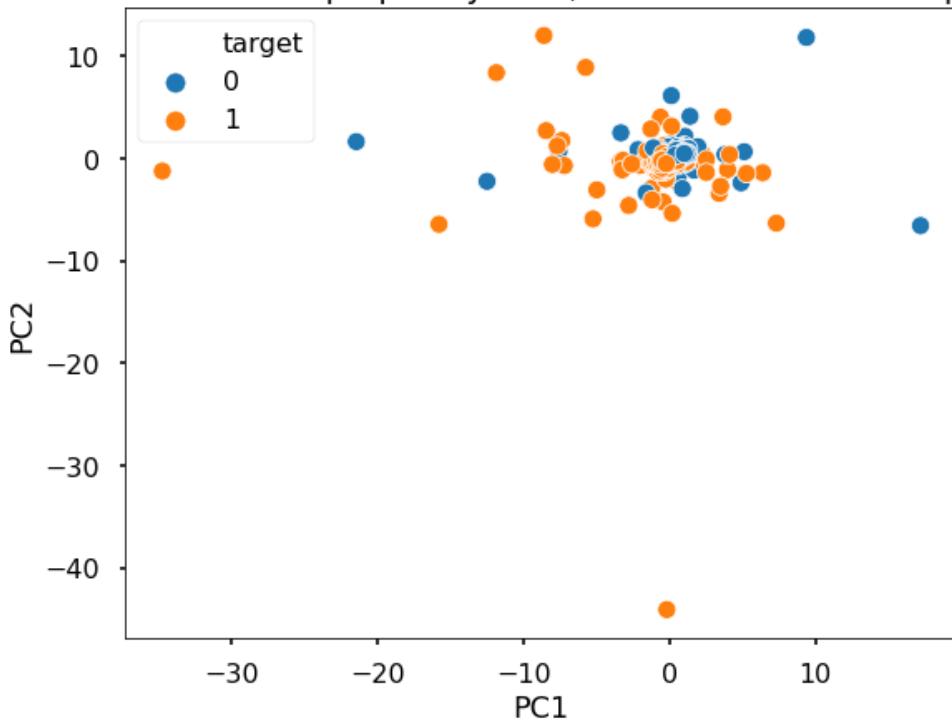
t-SNE visualization with perplexity -- 30, iterations -- 3000 and epsilon -- 100



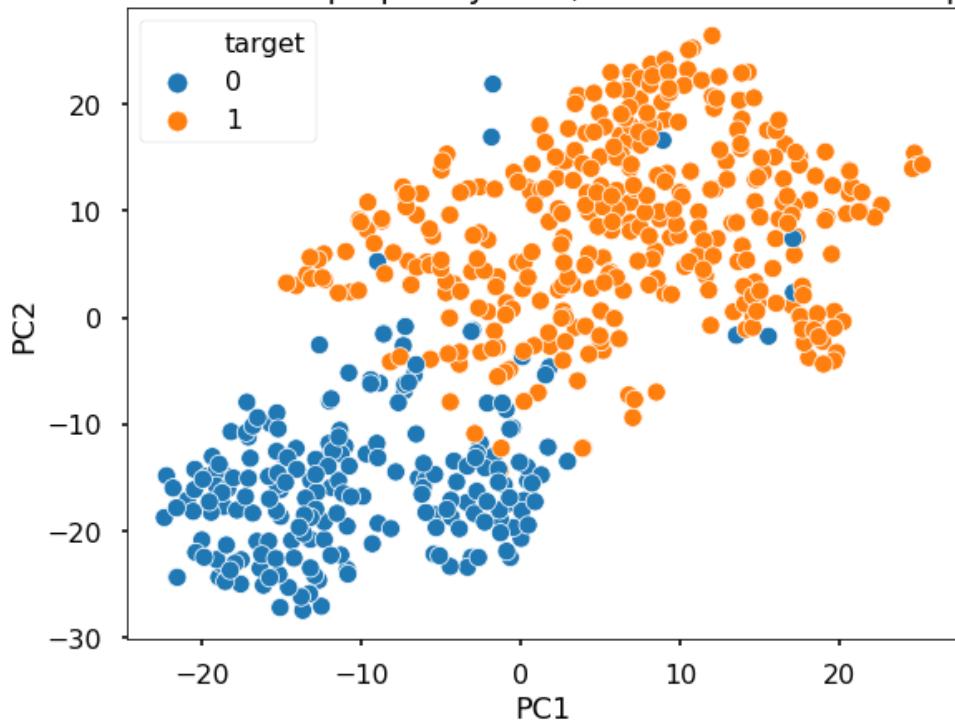
t-SNE visualization with perplexity -- 30, iterations -- 5000 and epsilon -- 100



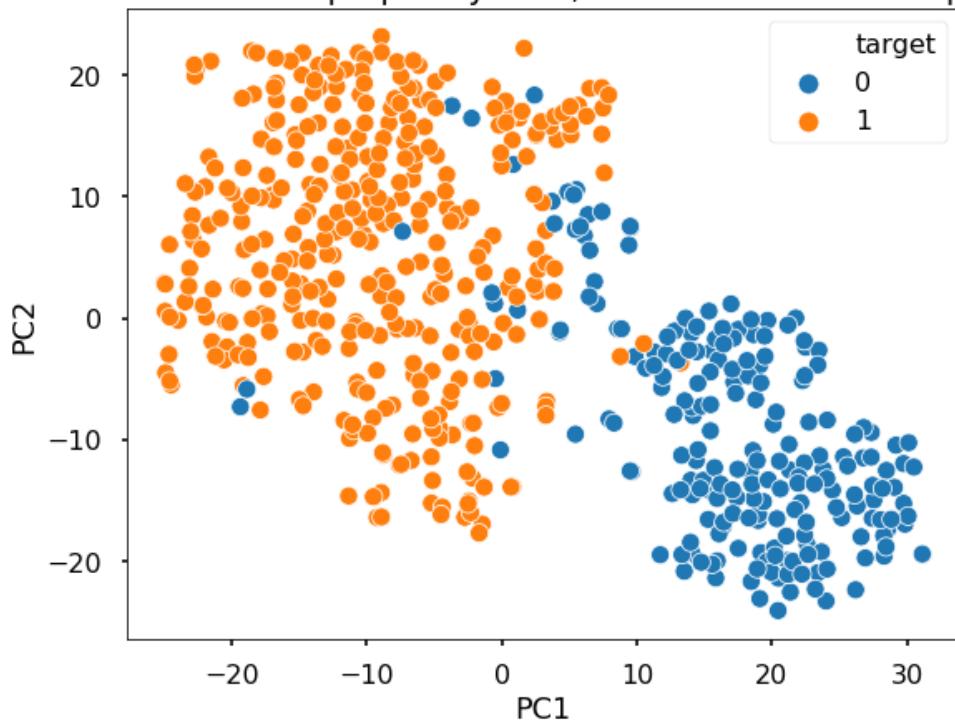
t-SNE visualization with perplexity -- 30, iterations -- 250 and epsilon -- 200



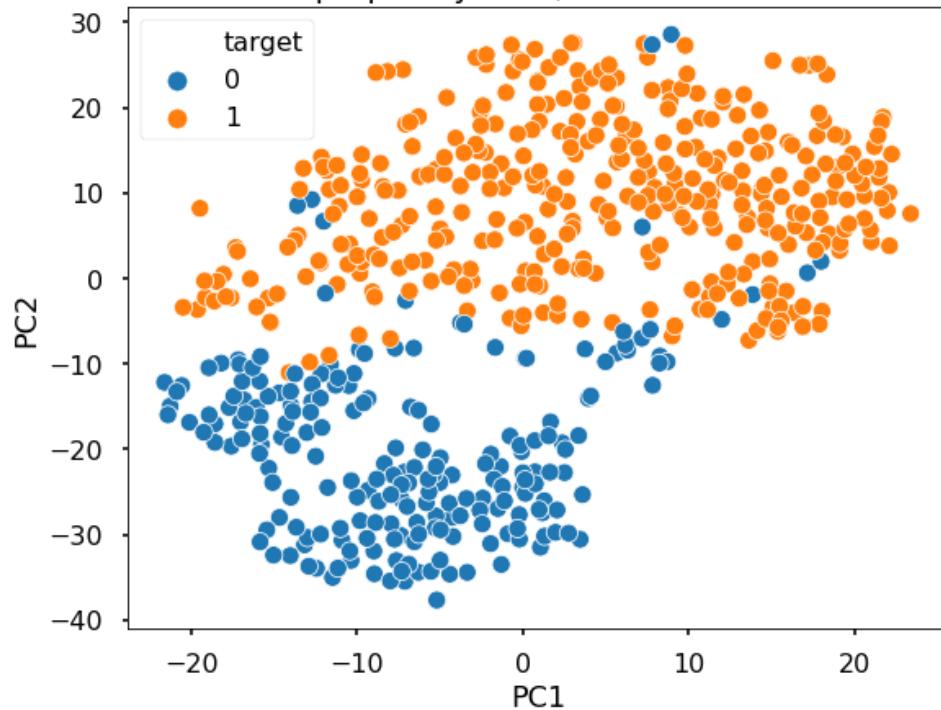
t-SNE visualization with perplexity -- 30, iterations -- 500 and epsilon -- 200



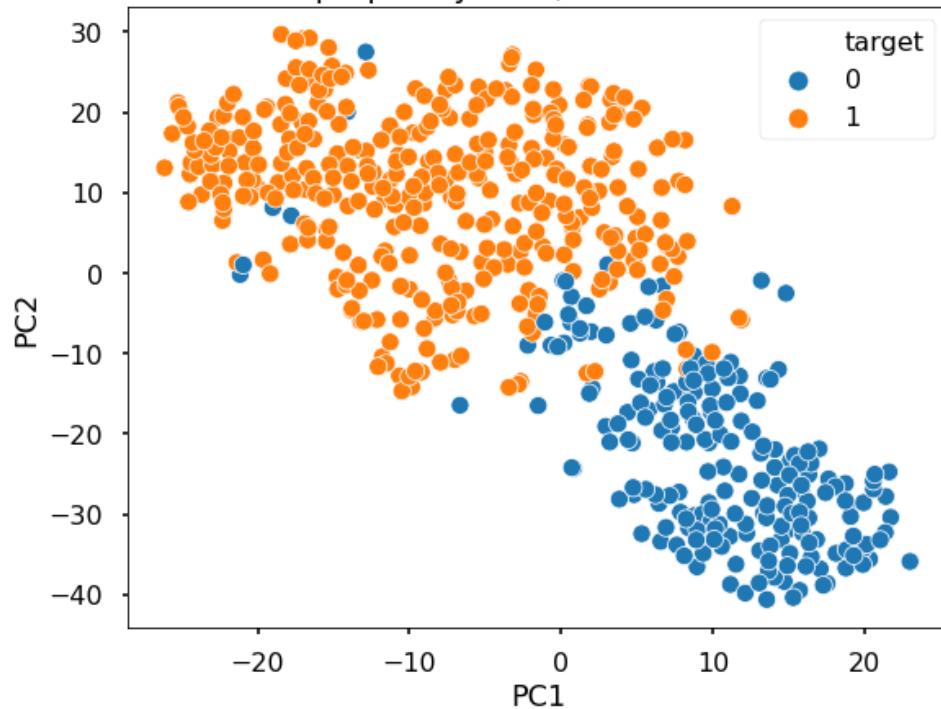
t-SNE visualization with perplexity -- 30, iterations -- 750 and epsilon -- 200



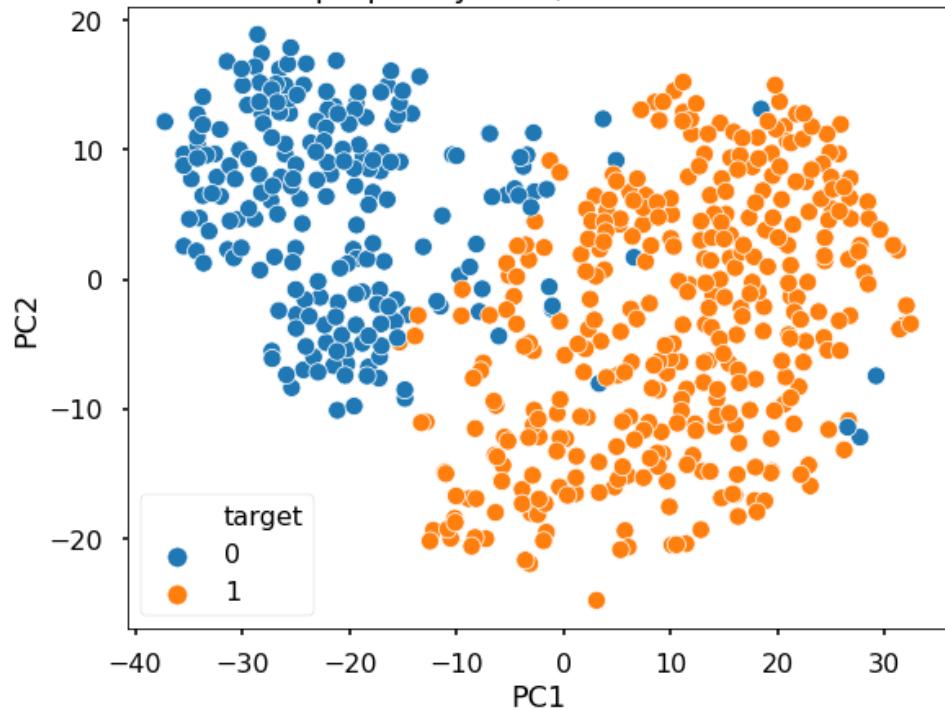
t-SNE visualization with perplexity -- 30, iterations -- 1000 and epsilon -- 200



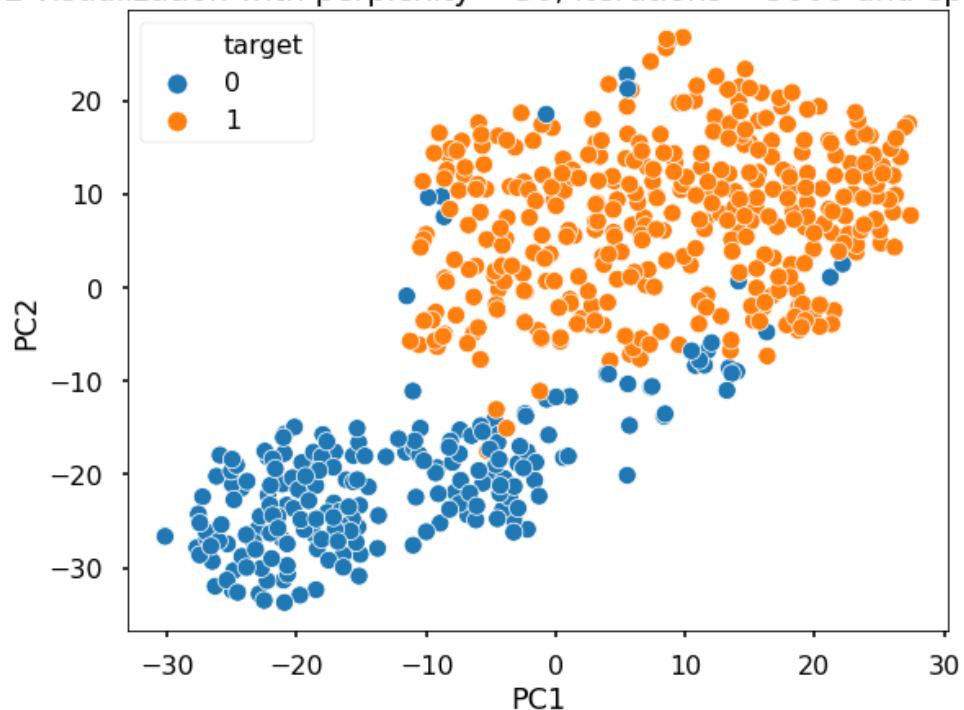
t-SNE visualization with perplexity -- 30, iterations -- 2000 and epsilon -- 200



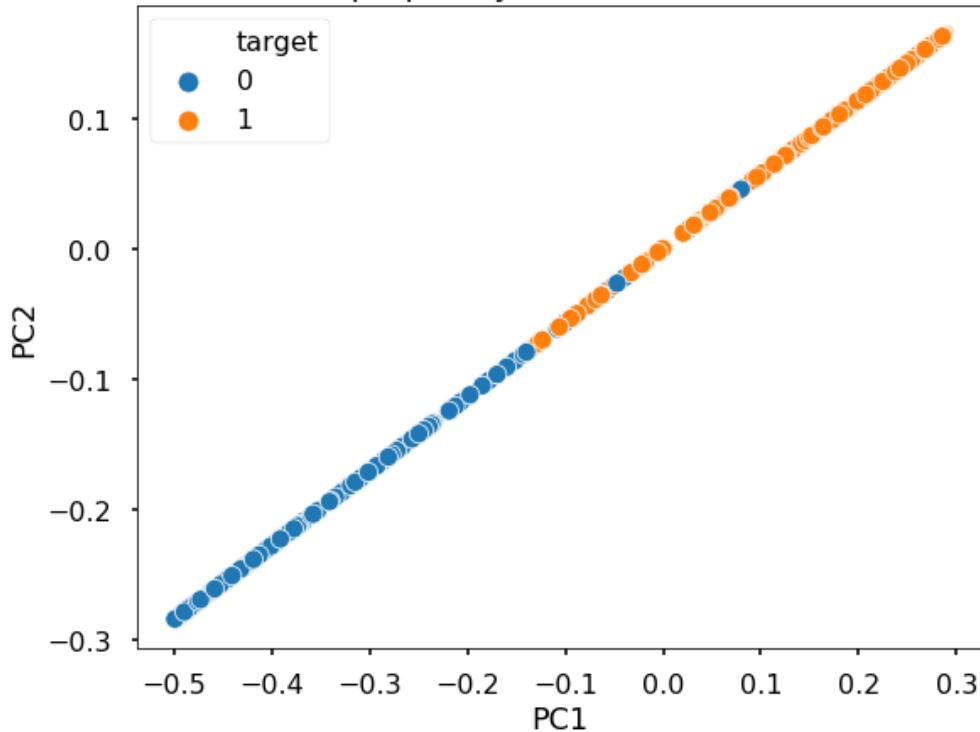
t-SNE visualization with perplexity -- 30, iterations -- 3000 and epsilon -- 200



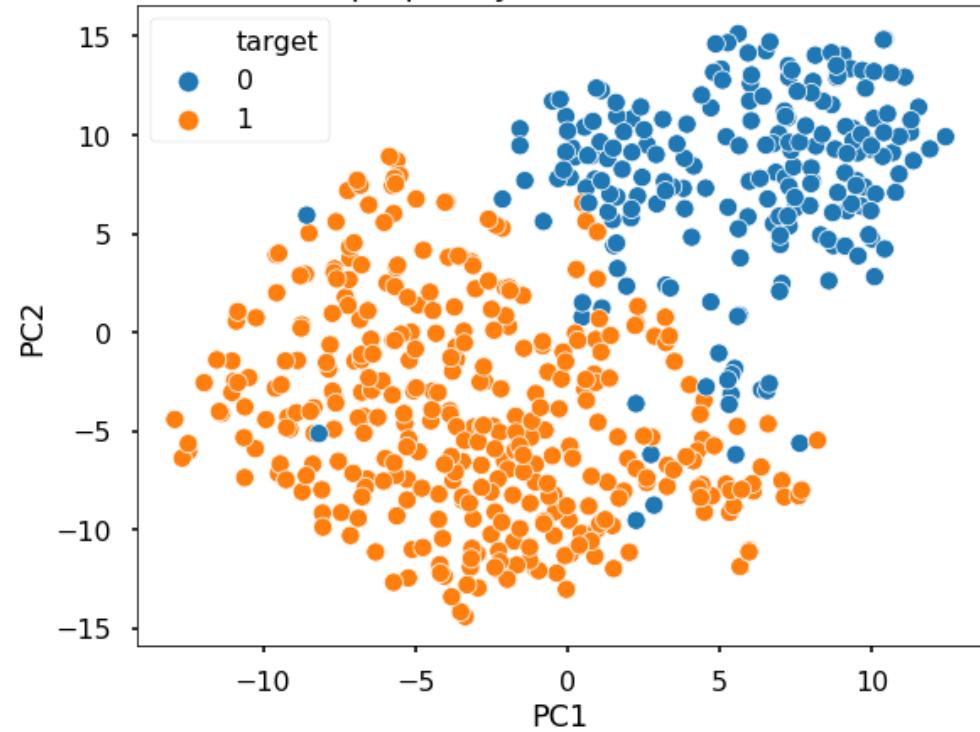
t-SNE visualization with perplexity -- 30, iterations -- 5000 and epsilon -- 200



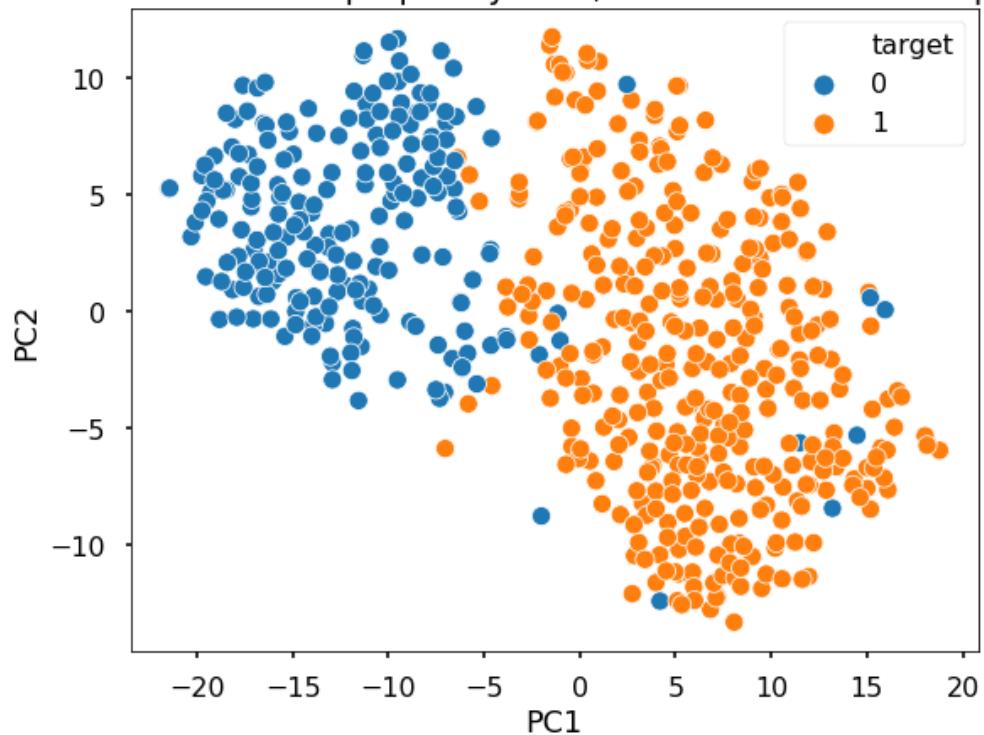
t-SNE visualization with perplexity -- 50, iterations -- 250 and epsilon -- 10



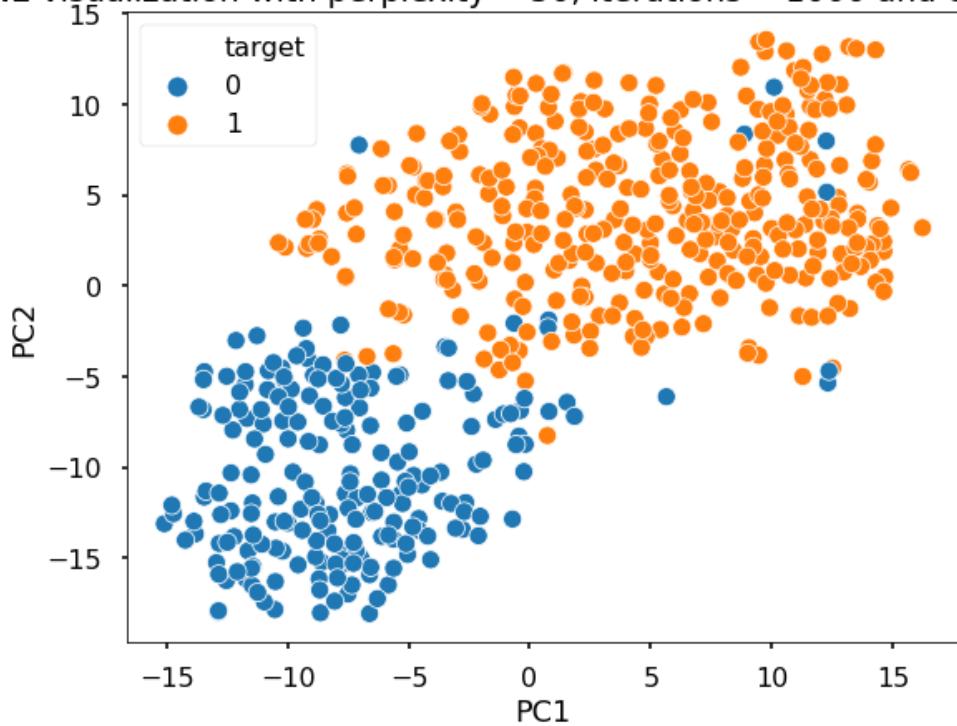
t-SNE visualization with perplexity -- 50, iterations -- 500 and epsilon -- 10



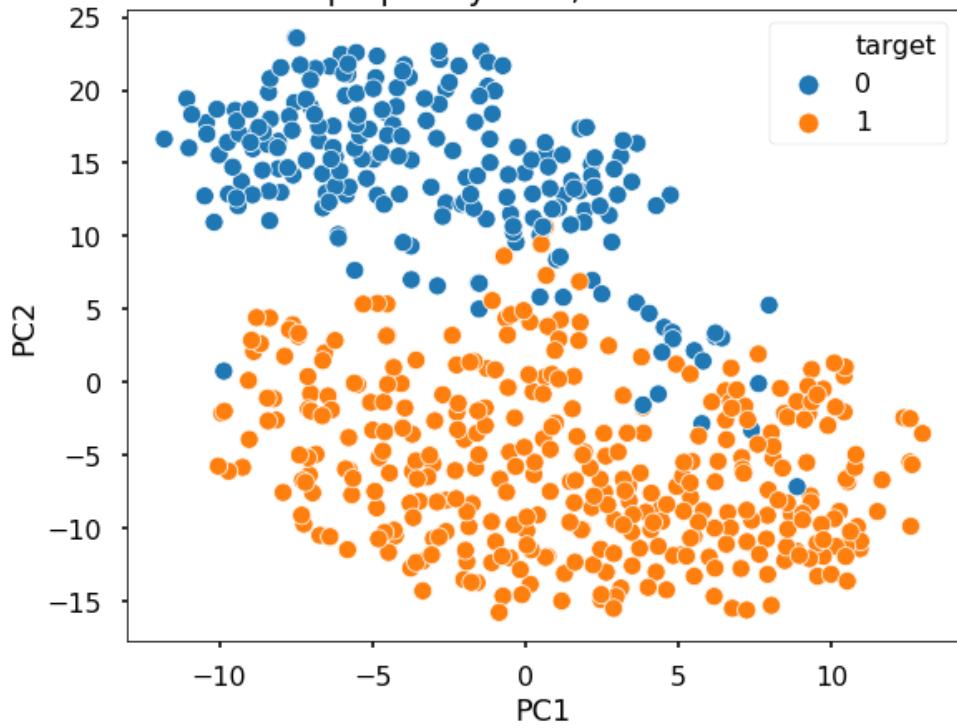
t-SNE visualization with perplexity -- 50, iterations -- 750 and epsilon -- 10



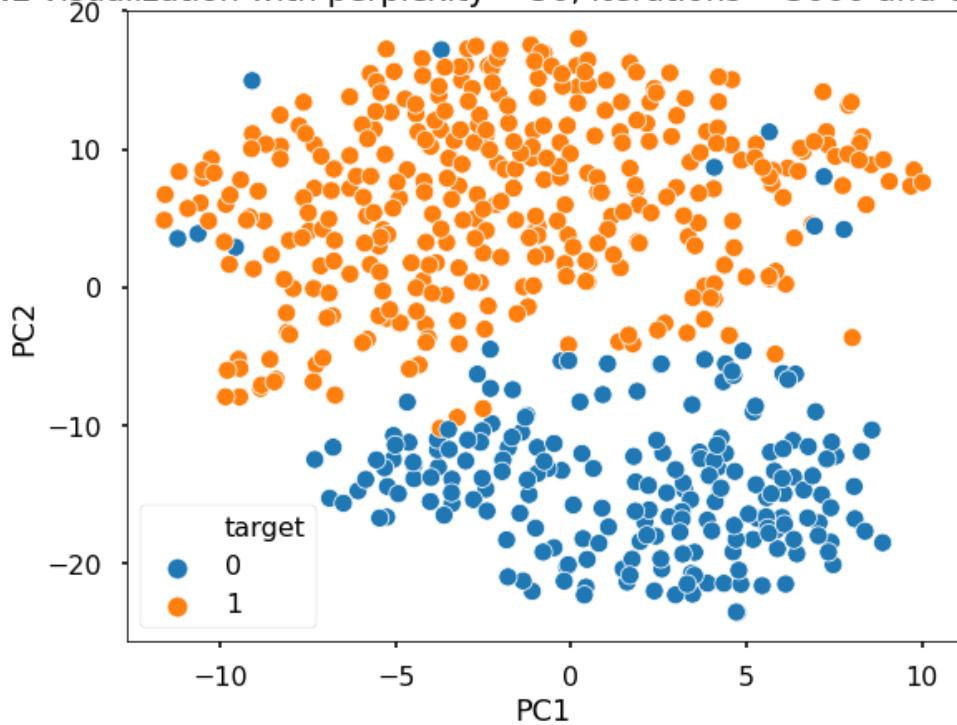
t-SNE visualization with perplexity -- 50, iterations -- 1000 and epsilon -- 10



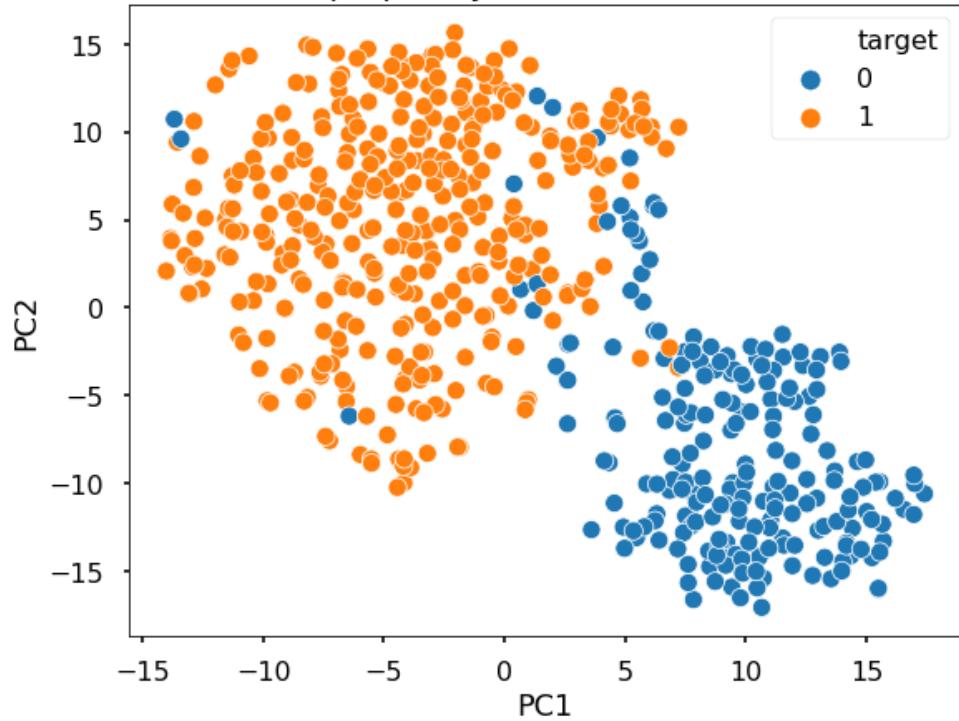
t-SNE visualization with perplexity -- 50, iterations -- 2000 and epsilon -- 10



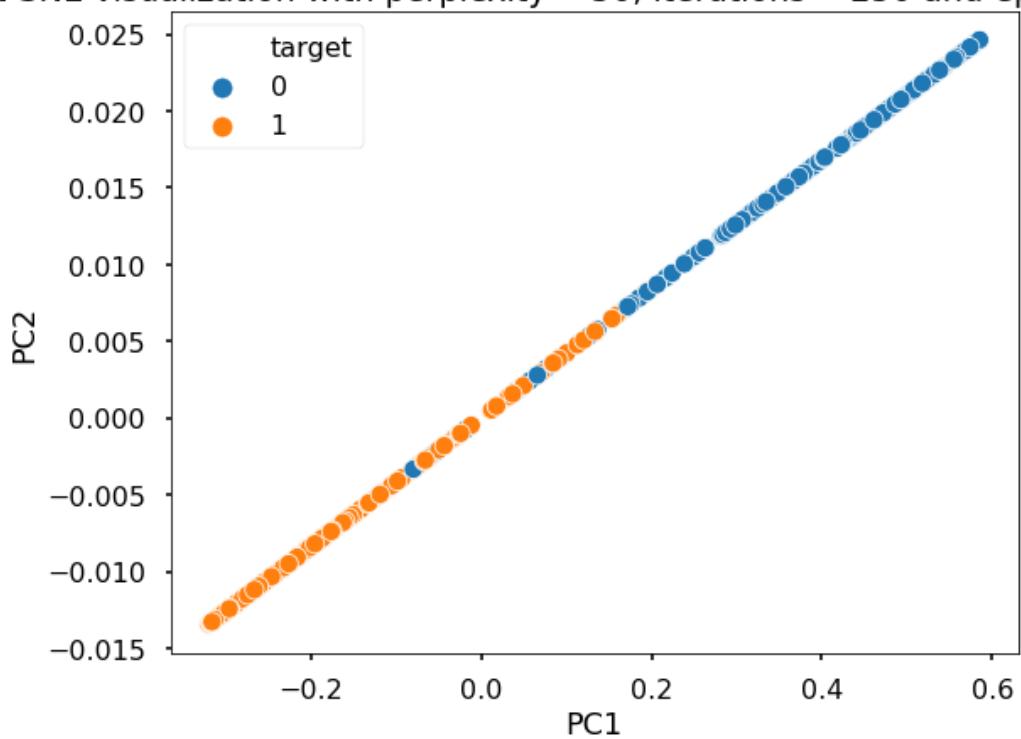
t-SNE visualization with perplexity -- 50, iterations -- 3000 and epsilon -- 10



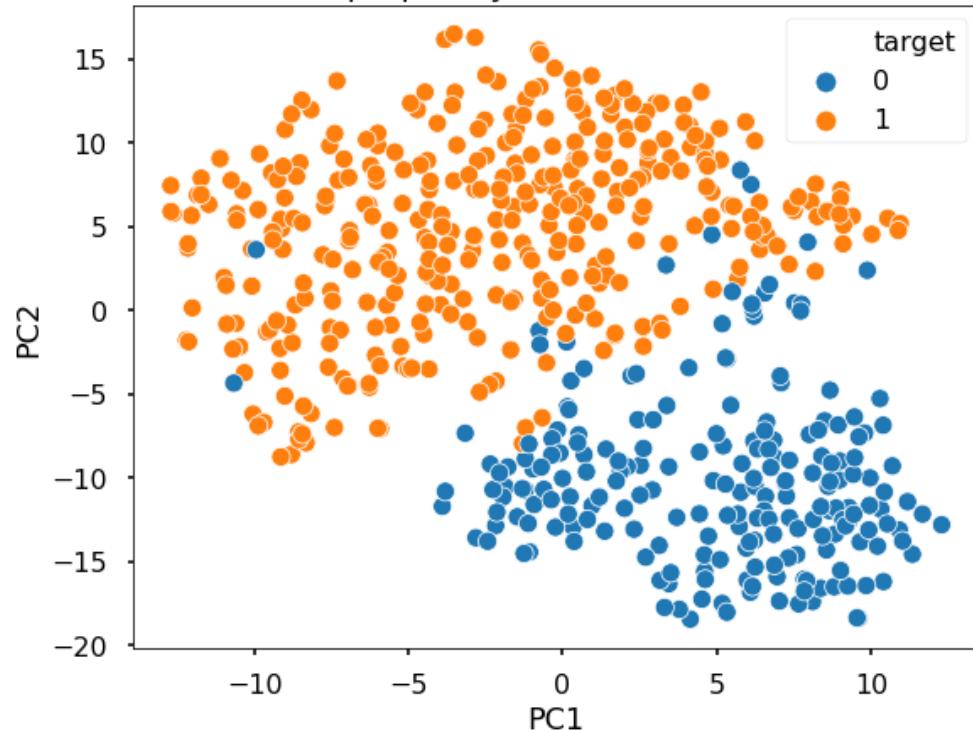
t-SNE visualization with perplexity -- 50, iterations -- 5000 and epsilon -- 10



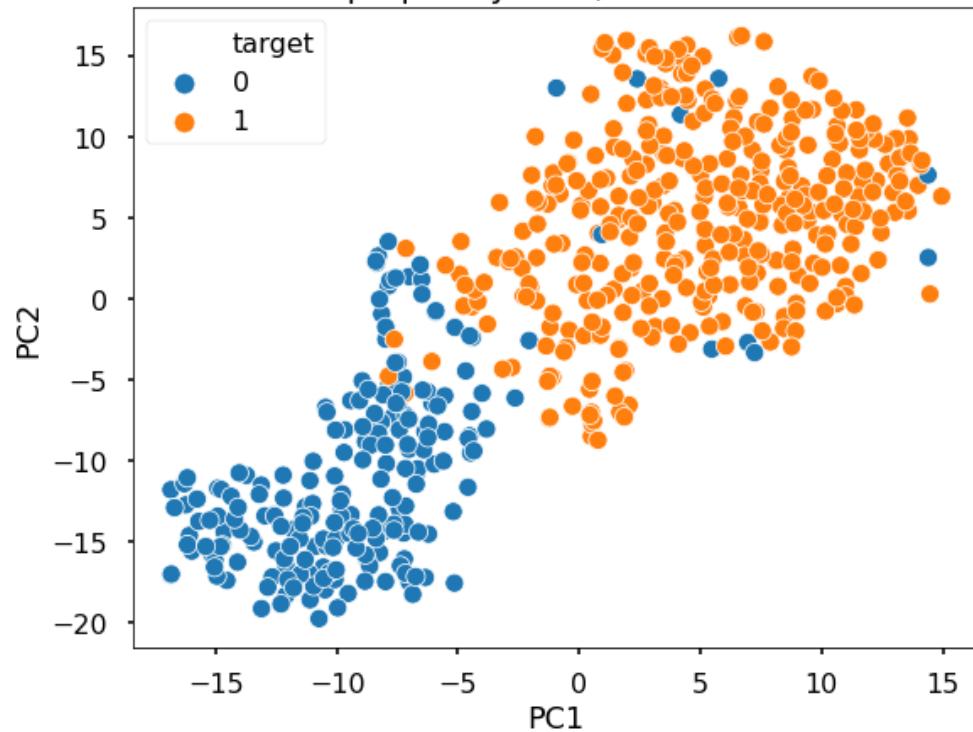
t-SNE visualization with perplexity -- 50, iterations -- 250 and epsilon -- 30



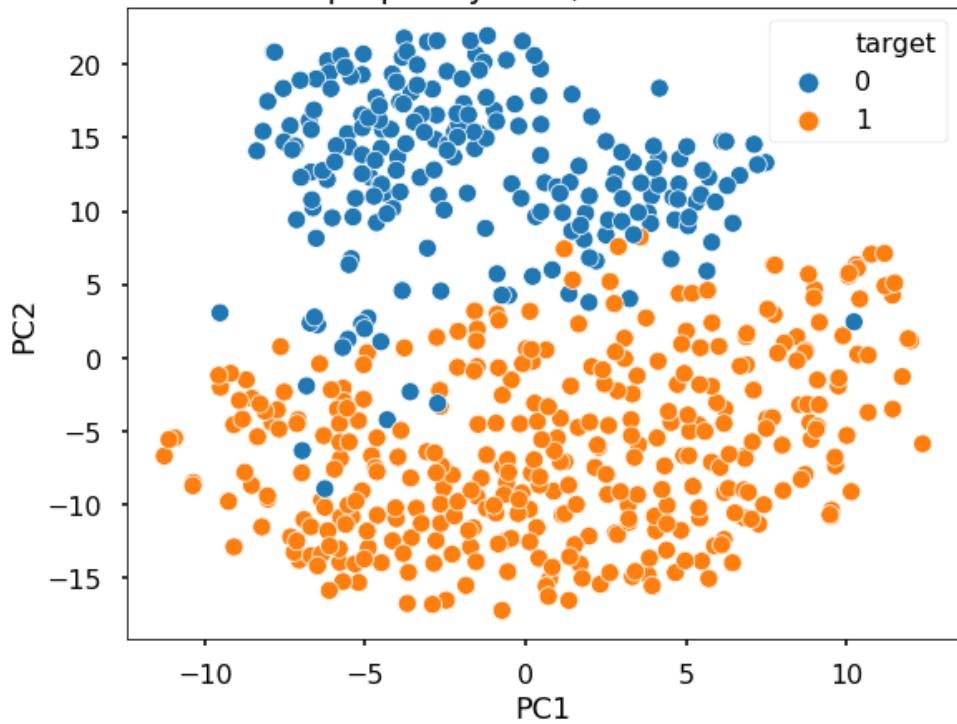
t-SNE visualization with perplexity -- 50, iterations -- 500 and epsilon -- 30



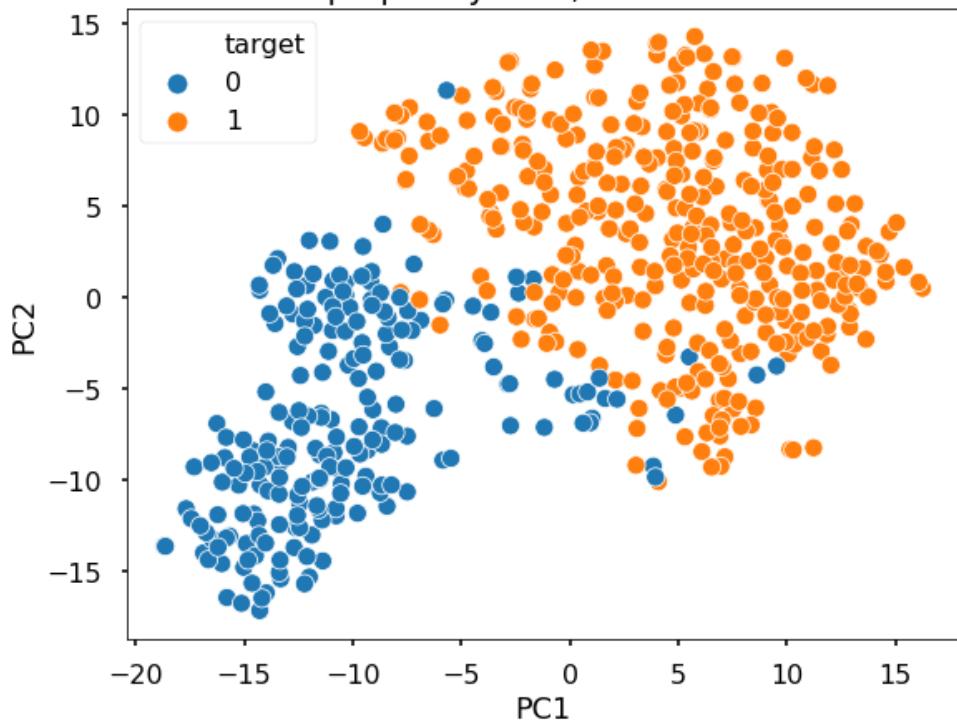
t-SNE visualization with perplexity -- 50, iterations -- 750 and epsilon -- 30



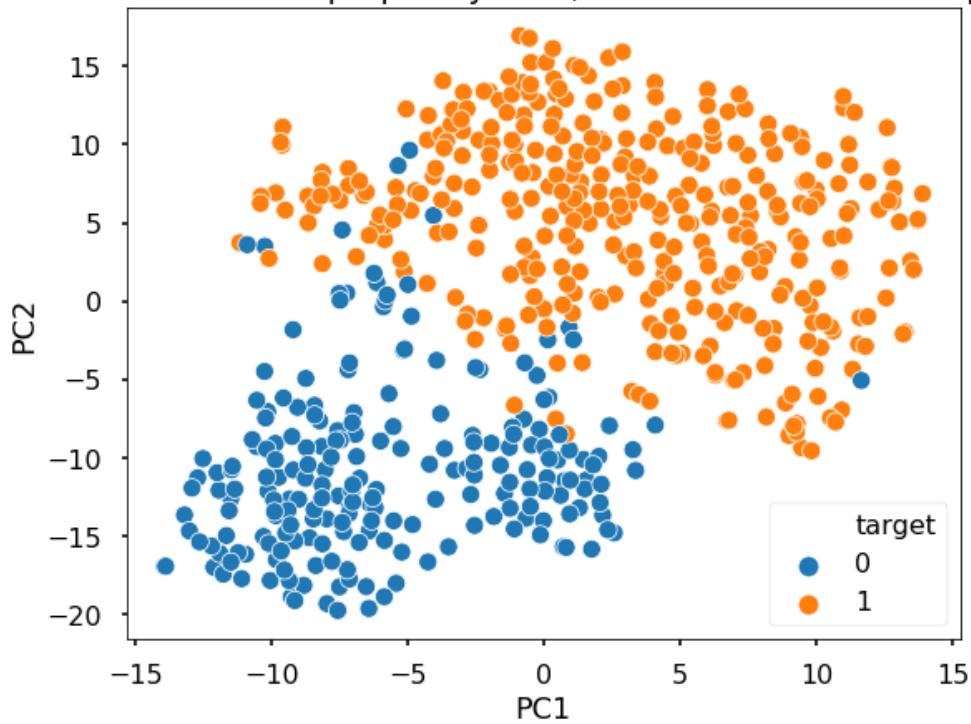
t-SNE visualization with perplexity -- 50, iterations -- 1000 and epsilon -- 30



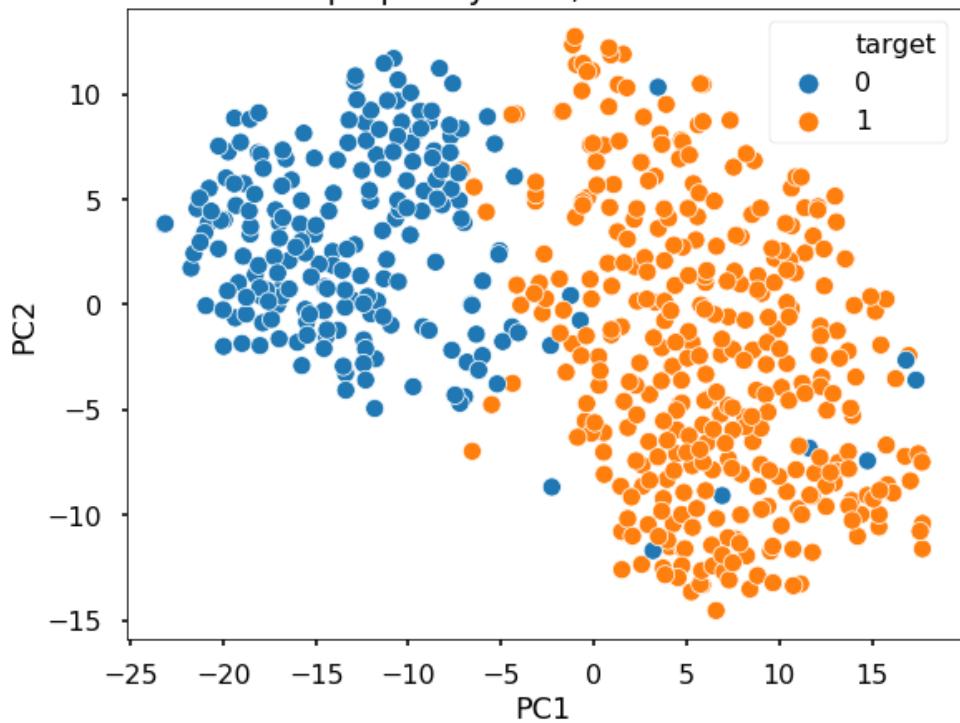
t-SNE visualization with perplexity -- 50, iterations -- 2000 and epsilon -- 30



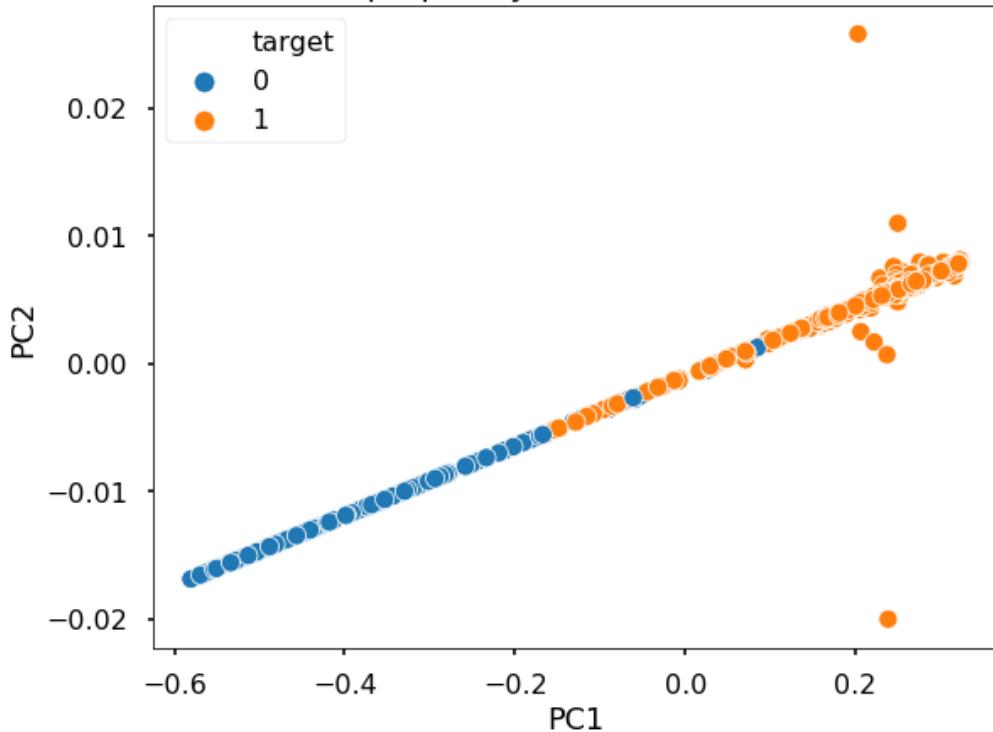
t-SNE visualization with perplexity -- 50, iterations -- 3000 and epsilon -- 30



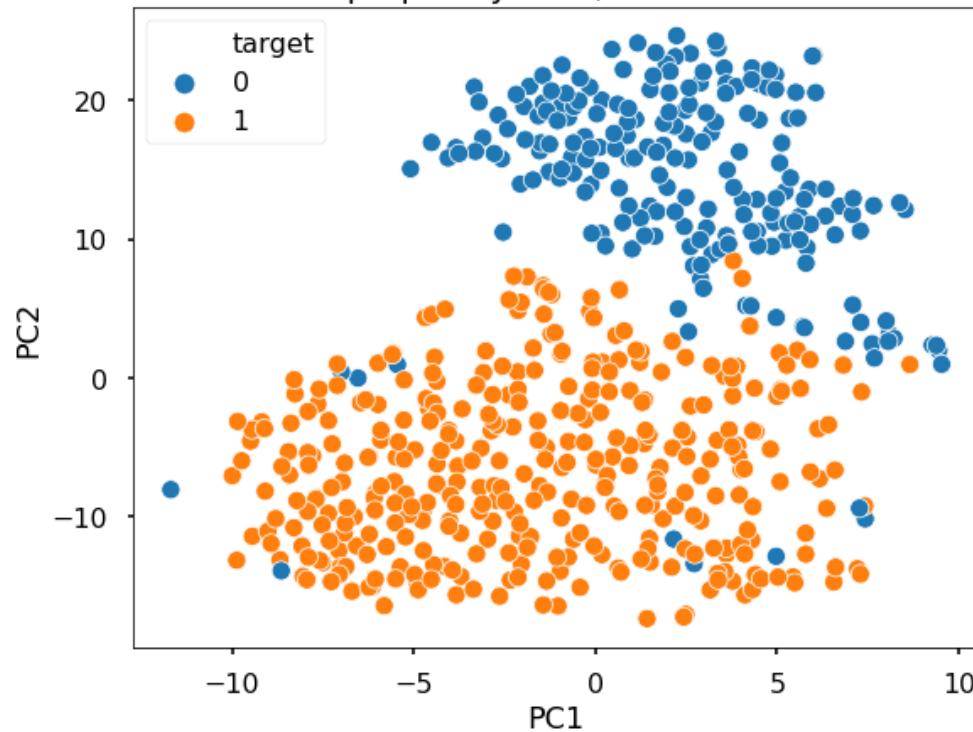
t-SNE visualization with perplexity -- 50, iterations -- 5000 and epsilon -- 30



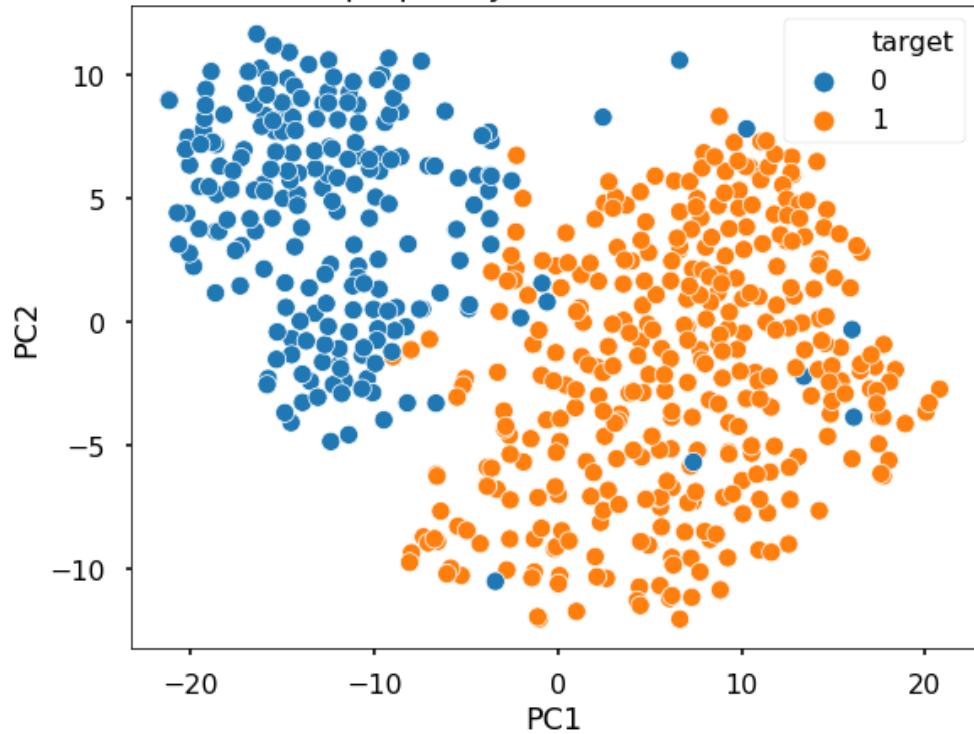
t-SNE visualization with perplexity -- 50, iterations -- 250 and epsilon -- 50



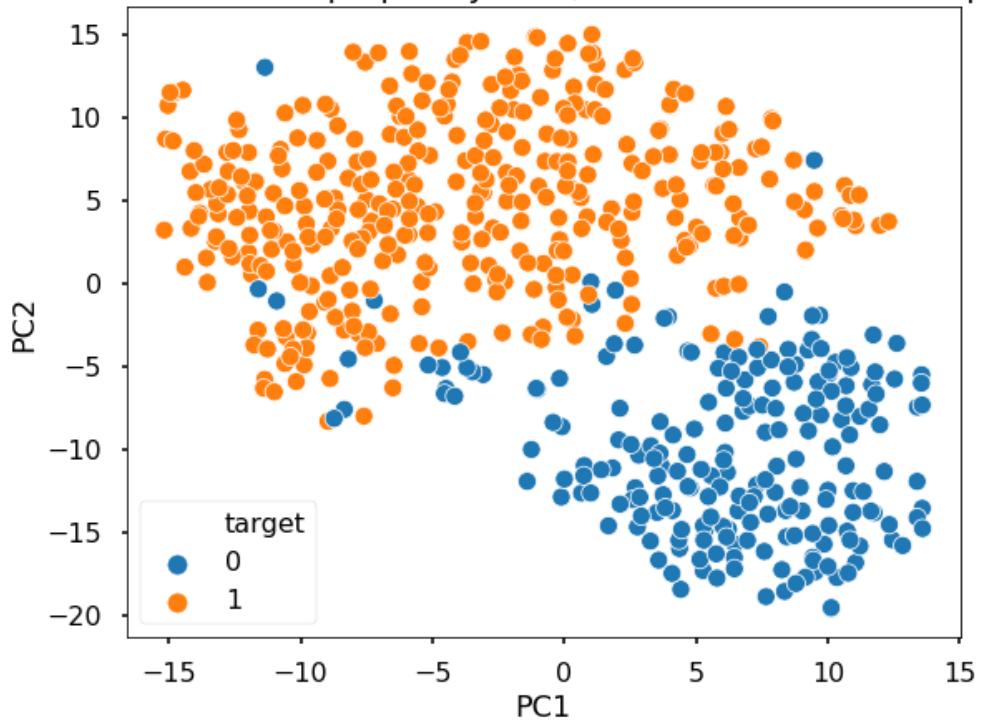
t-SNE visualization with perplexity -- 50, iterations -- 500 and epsilon -- 50



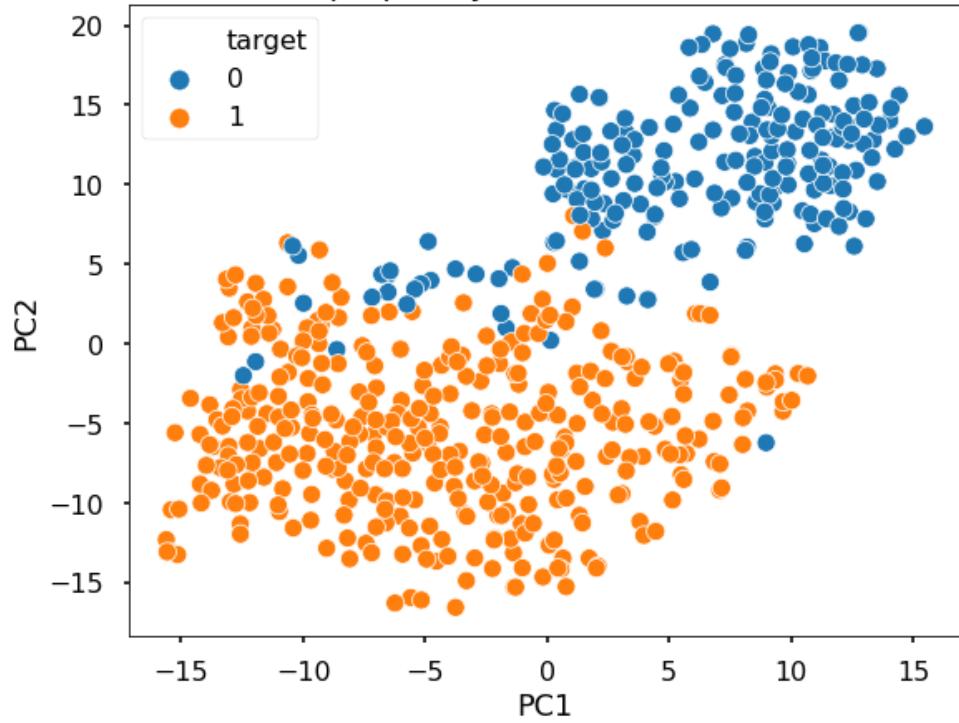
t-SNE visualization with perplexity -- 50, iterations -- 750 and epsilon -- 50



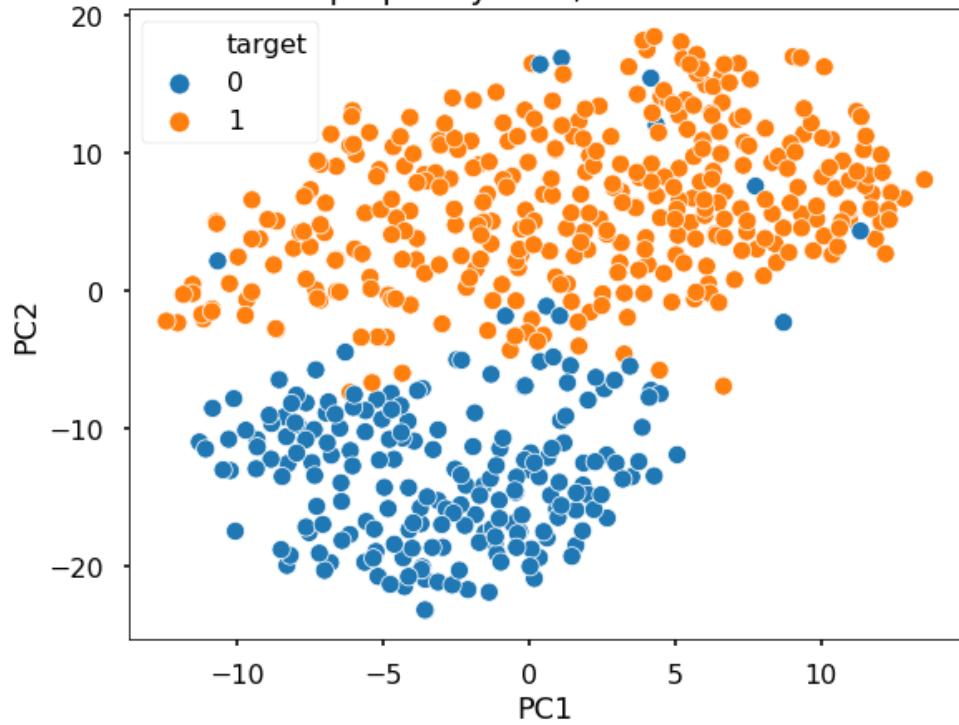
t-SNE visualization with perplexity -- 50, iterations -- 1000 and epsilon -- 50



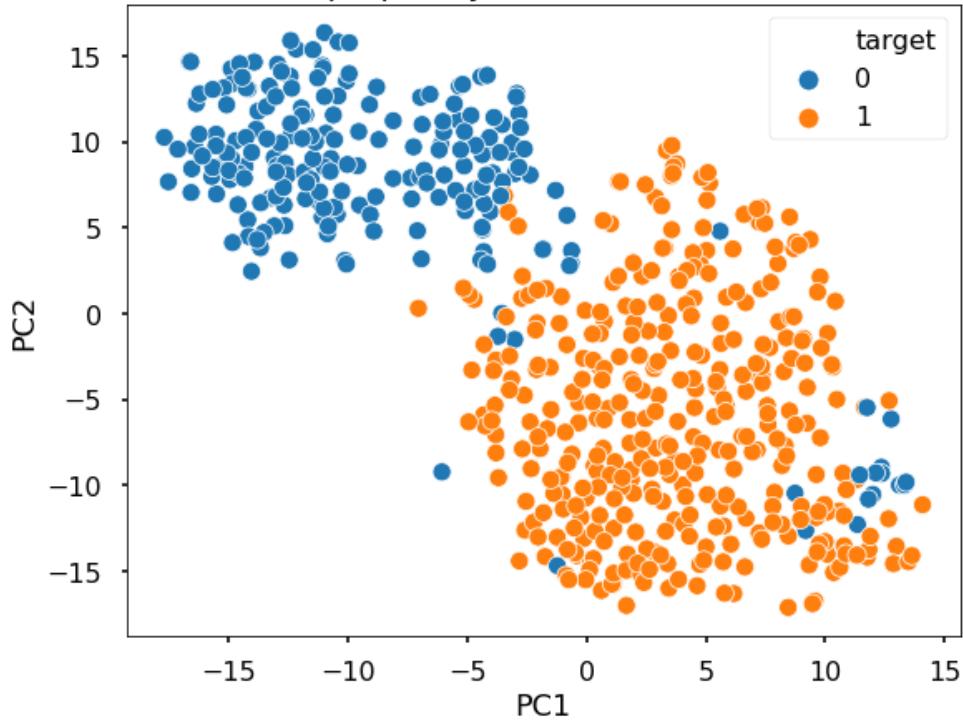
t-SNE visualization with perplexity -- 50, iterations -- 2000 and epsilon -- 50



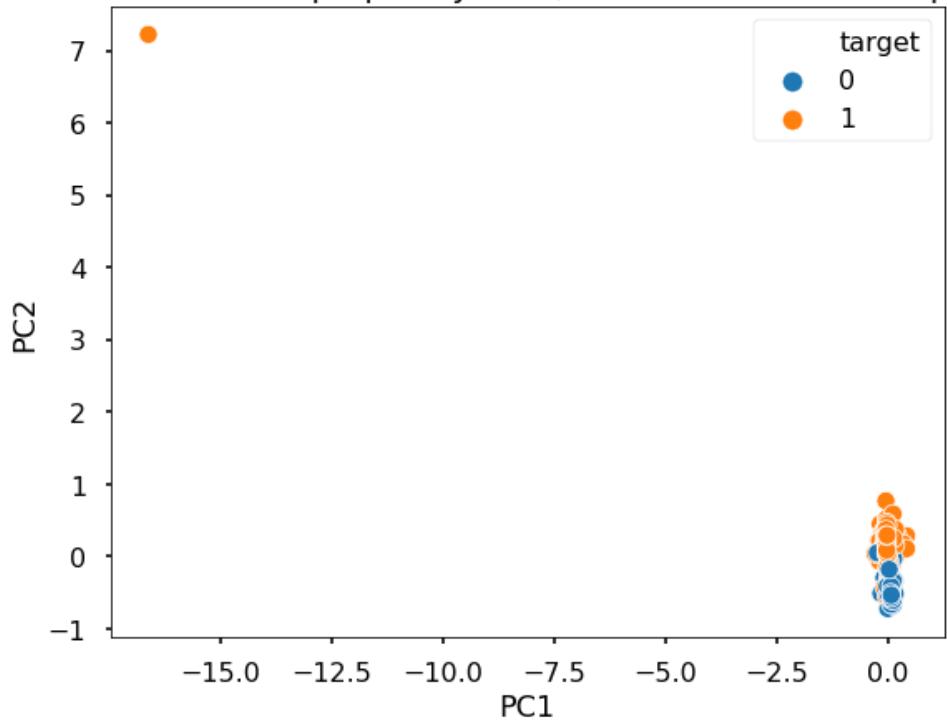
t-SNE visualization with perplexity -- 50, iterations -- 3000 and epsilon -- 50



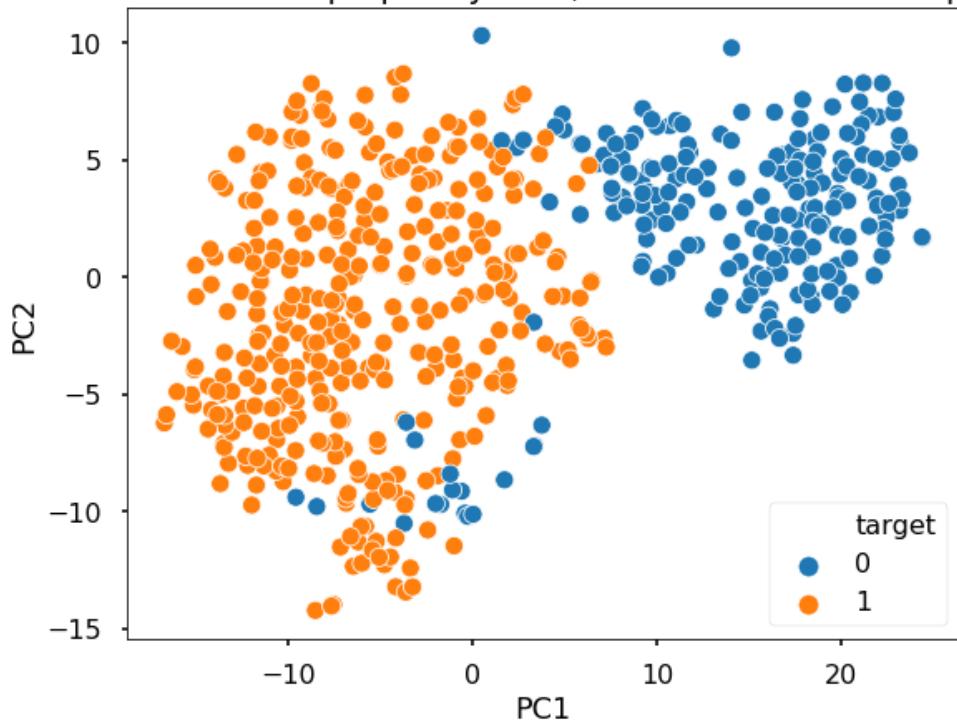
t-SNE visualization with perplexity -- 50, iterations -- 5000 and epsilon -- 50



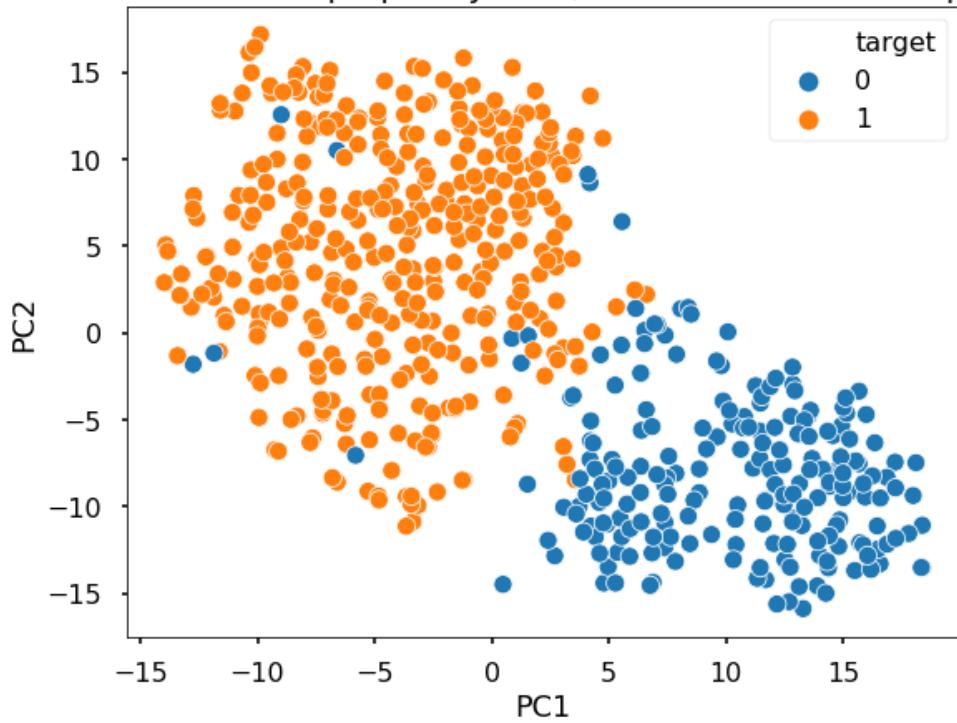
t-SNE visualization with perplexity -- 50, iterations -- 250 and epsilon -- 100



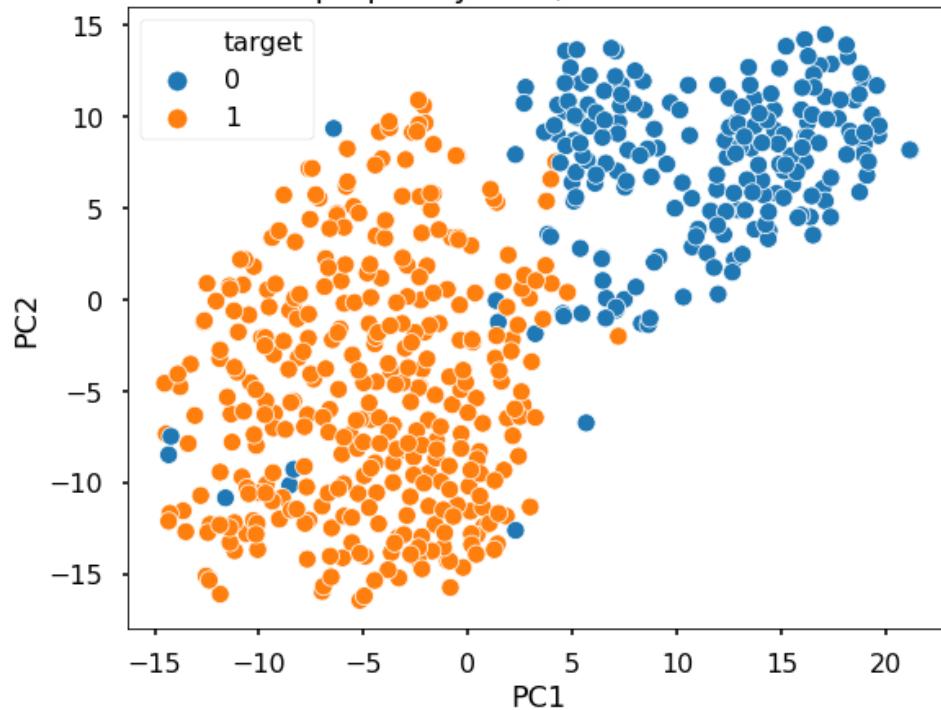
t-SNE visualization with perplexity -- 50, iterations -- 500 and epsilon -- 100



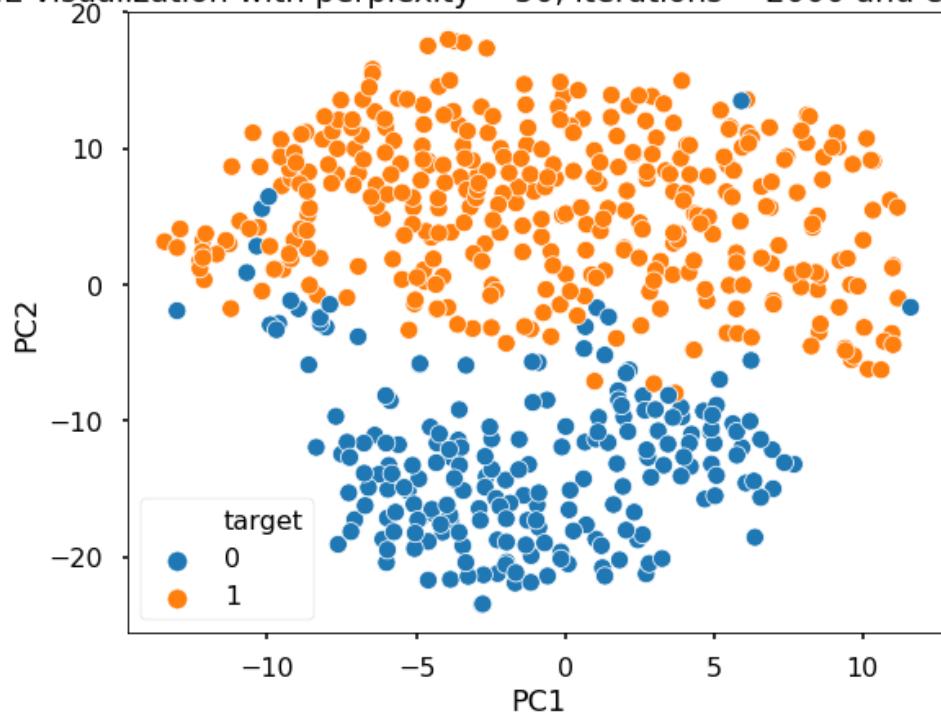
t-SNE visualization with perplexity -- 50, iterations -- 750 and epsilon -- 100



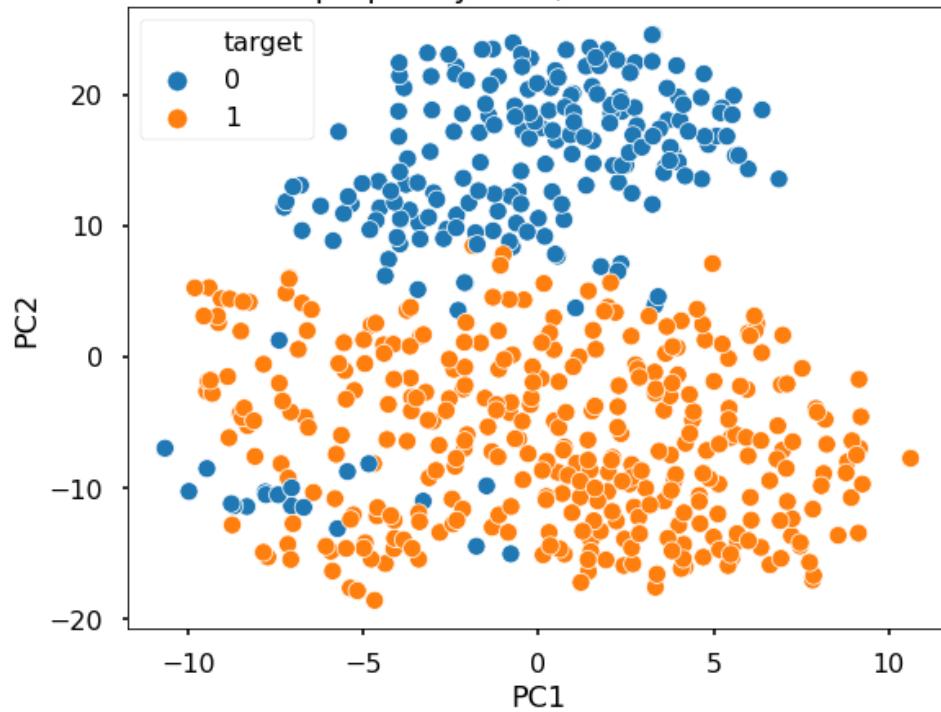
t-SNE visualization with perplexity -- 50, iterations -- 1000 and epsilon -- 100



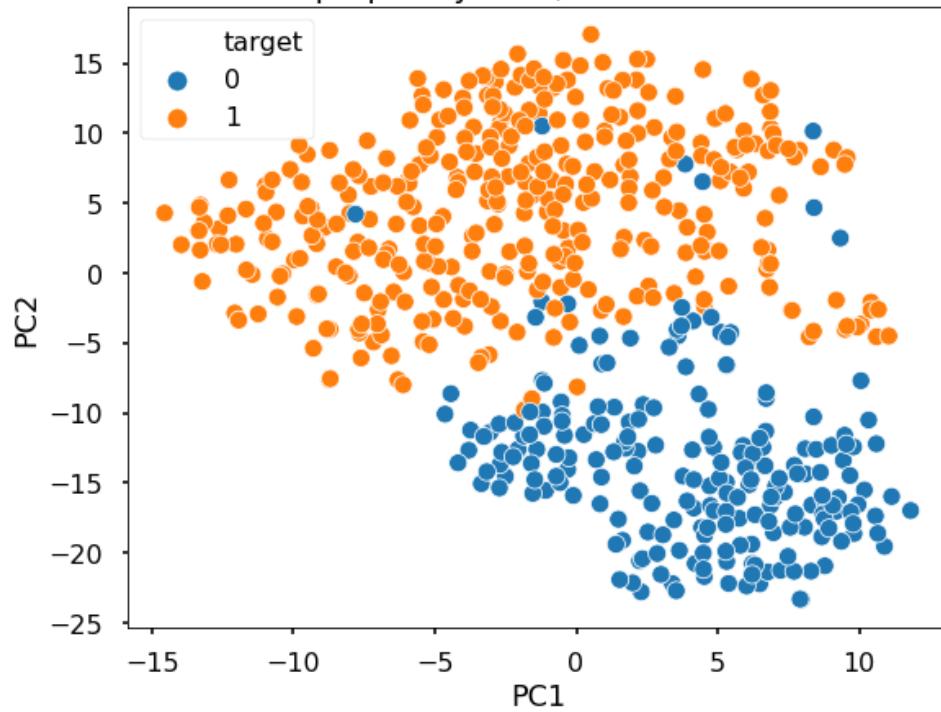
t-SNE visualization with perplexity -- 50, iterations -- 2000 and epsilon -- 100



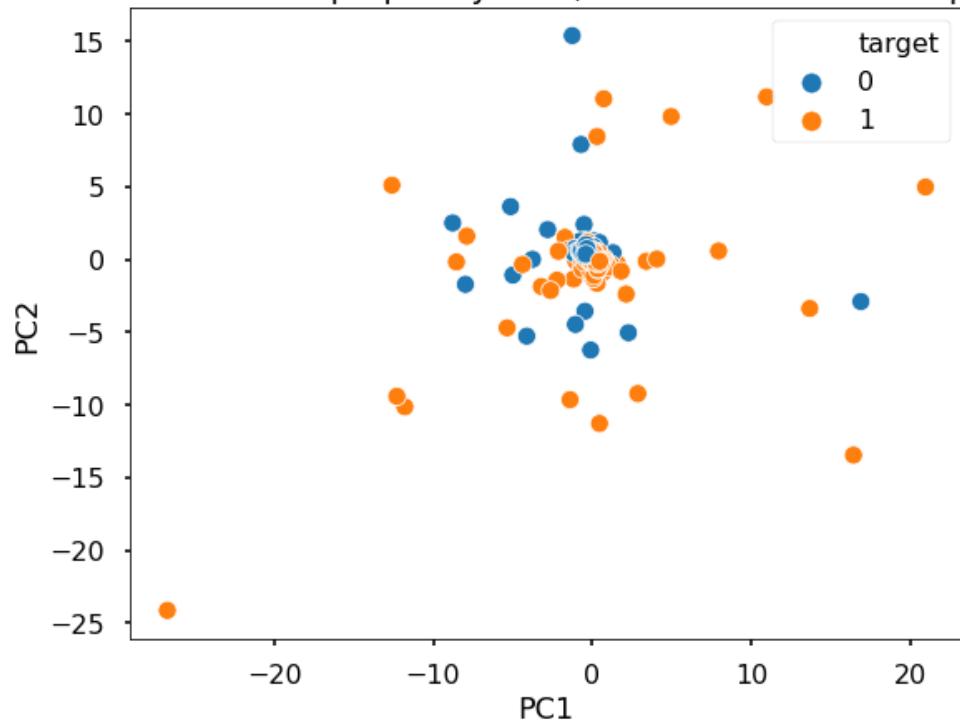
t-SNE visualization with perplexity -- 50, iterations -- 3000 and epsilon -- 100



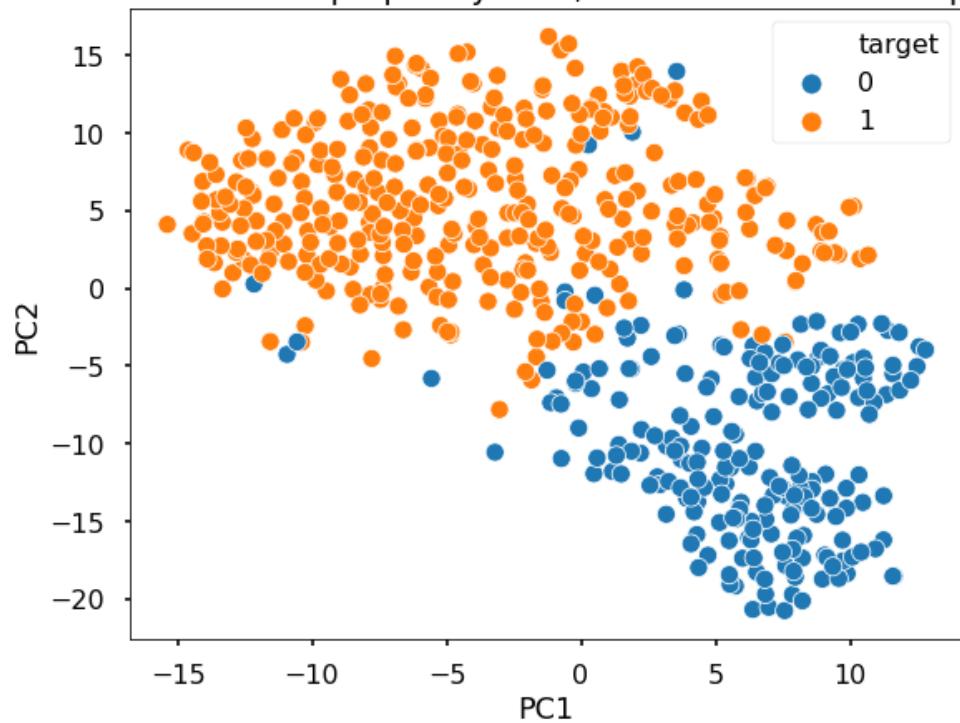
t-SNE visualization with perplexity -- 50, iterations -- 5000 and epsilon -- 100



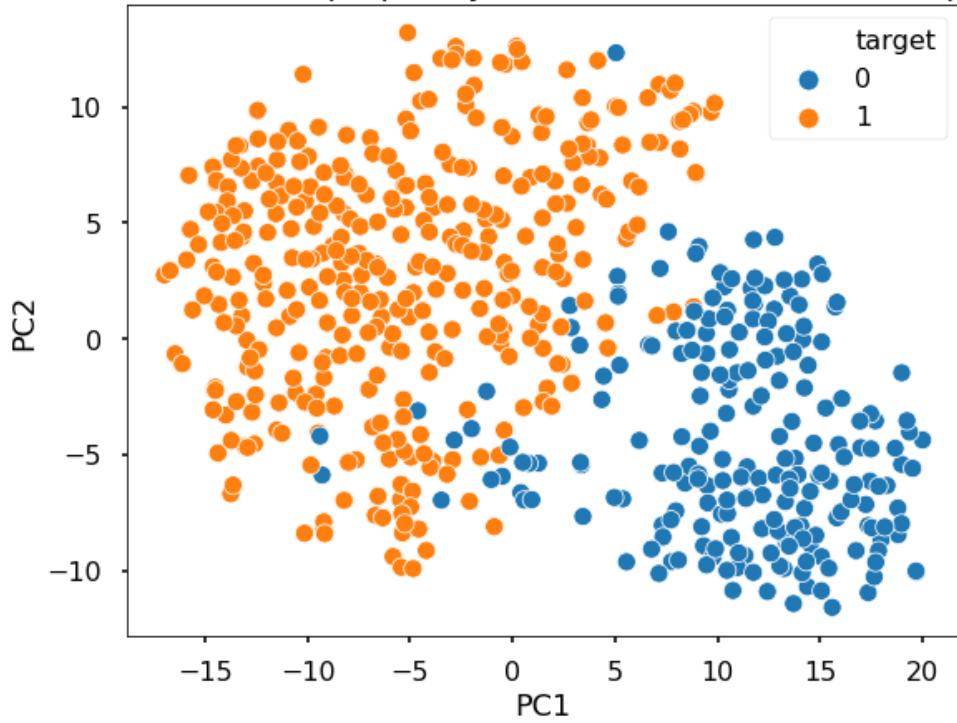
t-SNE visualization with perplexity -- 50, iterations -- 250 and epsilon -- 200



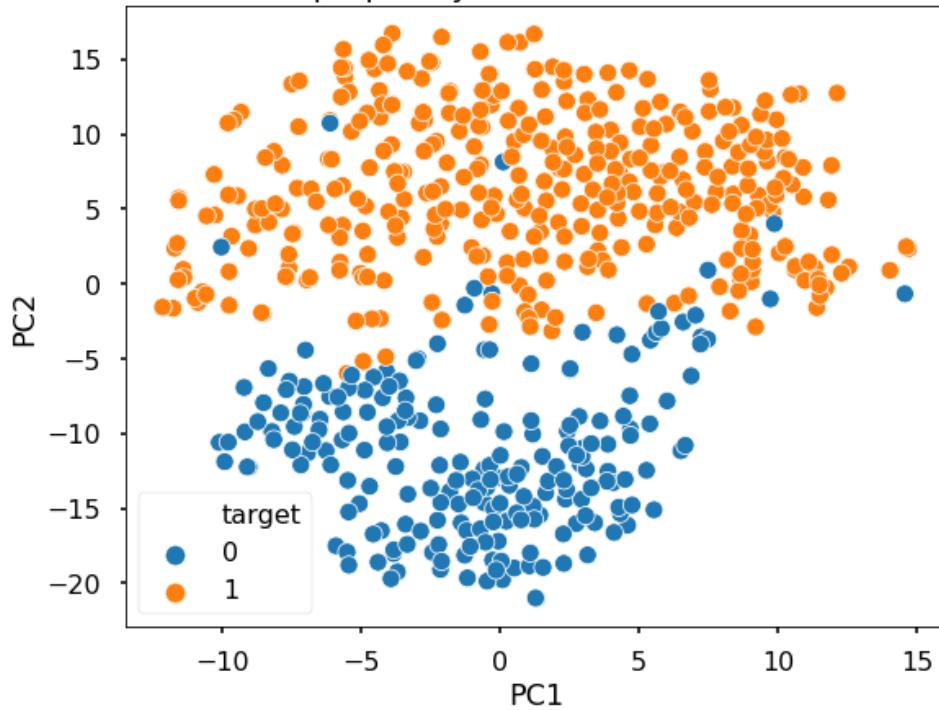
t-SNE visualization with perplexity -- 50, iterations -- 500 and epsilon -- 200



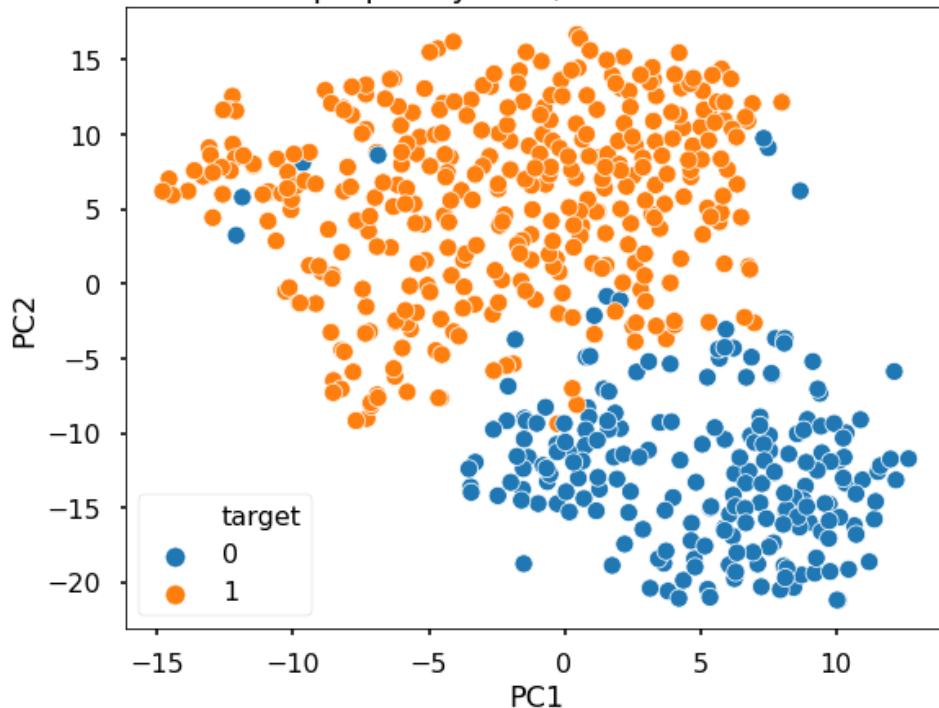
t-SNE visualization with perplexity -- 50, iterations -- 750 and epsilon -- 200



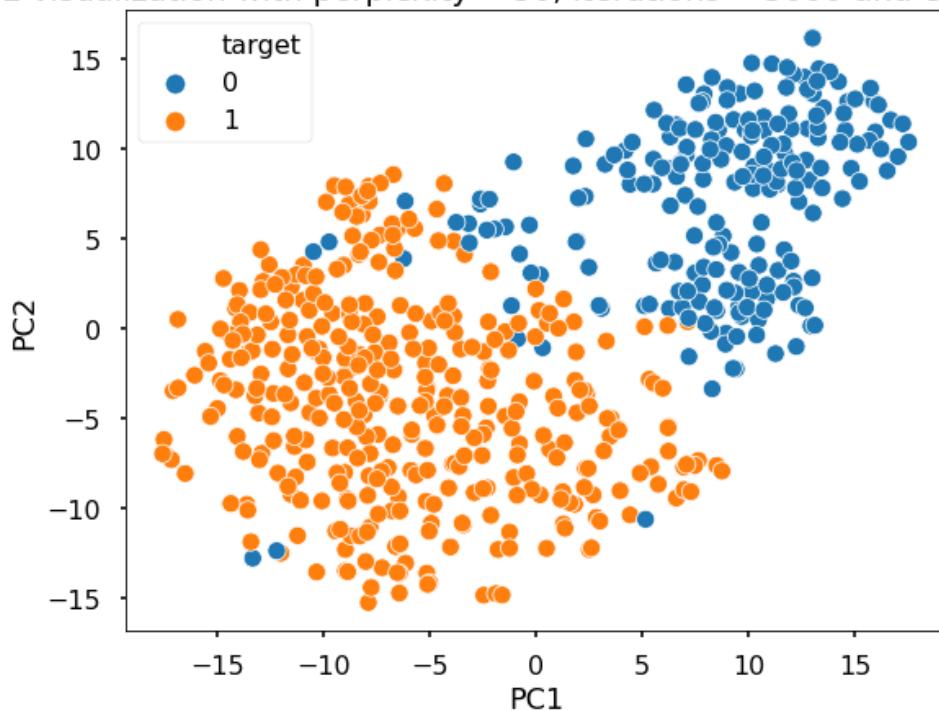
t-SNE visualization with perplexity -- 50, iterations -- 1000 and epsilon -- 200



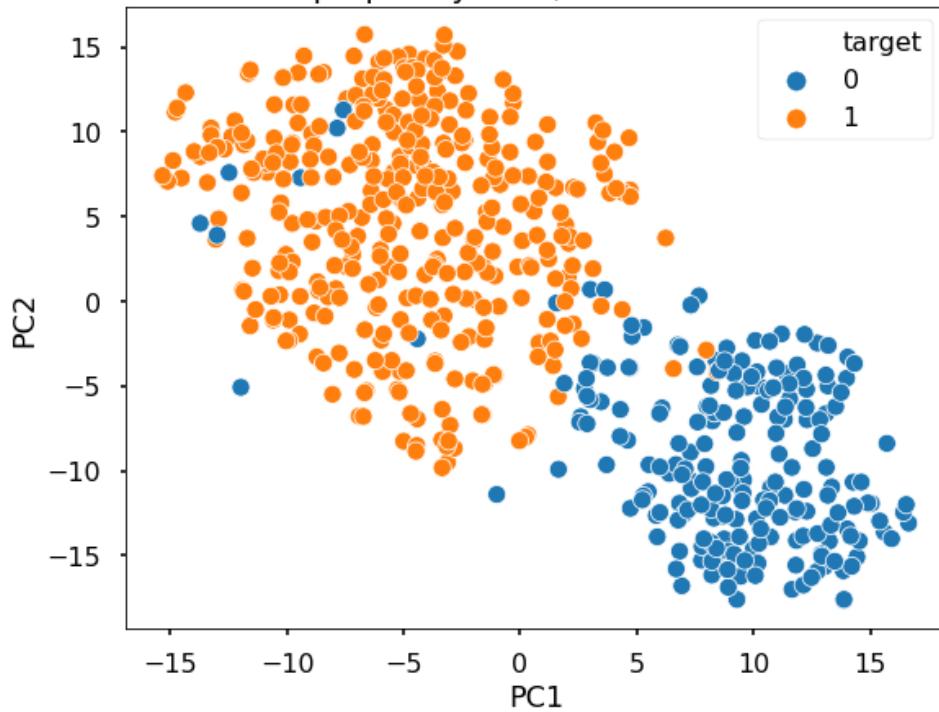
t-SNE visualization with perplexity -- 50, iterations -- 2000 and epsilon -- 200



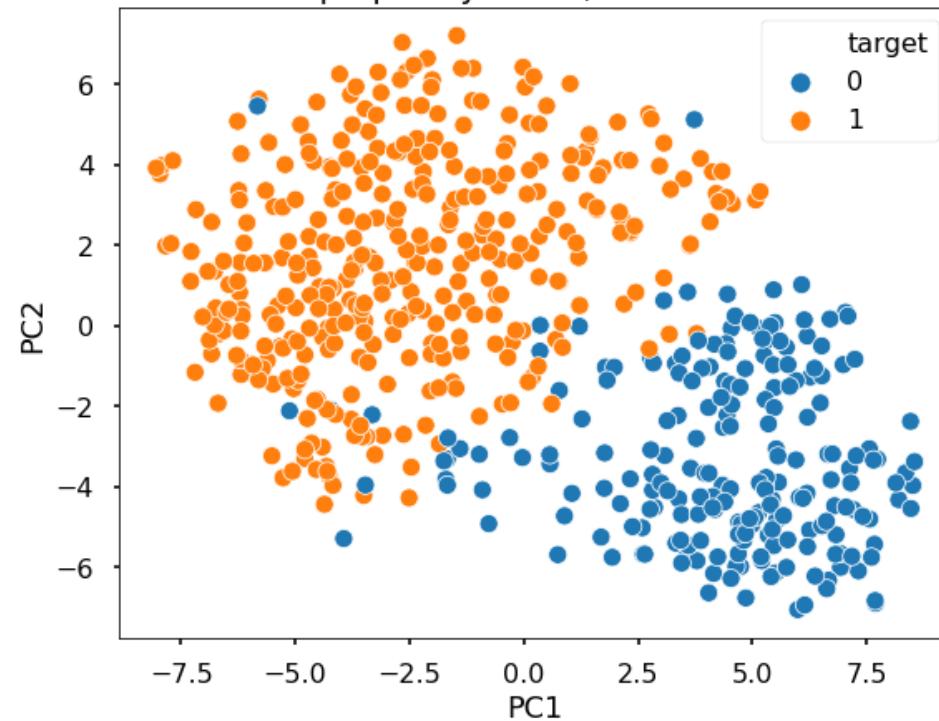
t-SNE visualization with perplexity -- 50, iterations -- 3000 and epsilon -- 200



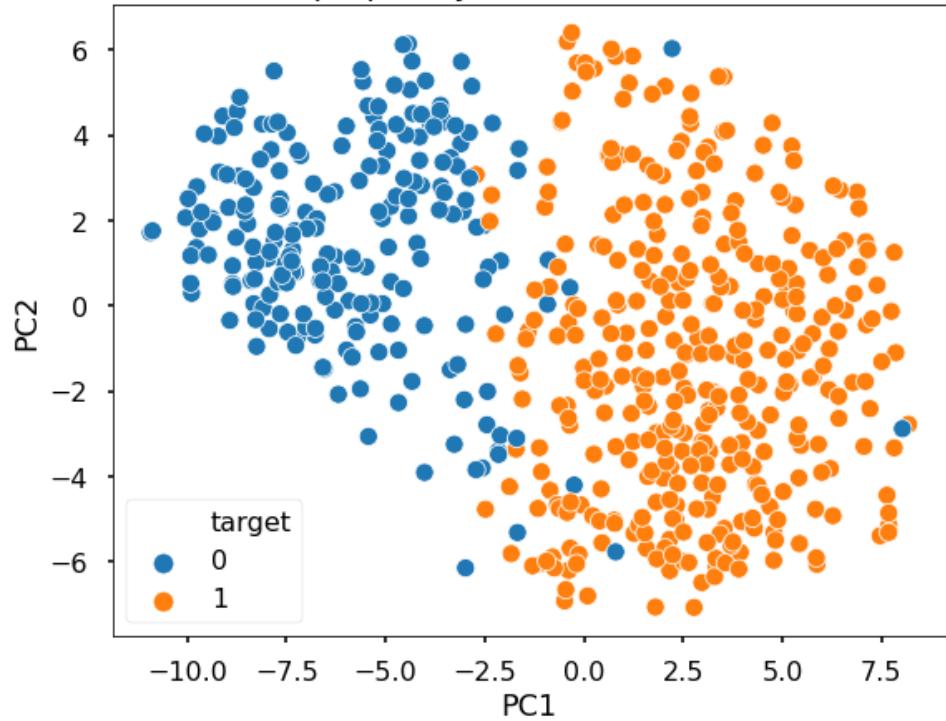
t-SNE visualization with perplexity -- 50, iterations -- 5000 and epsilon -- 200



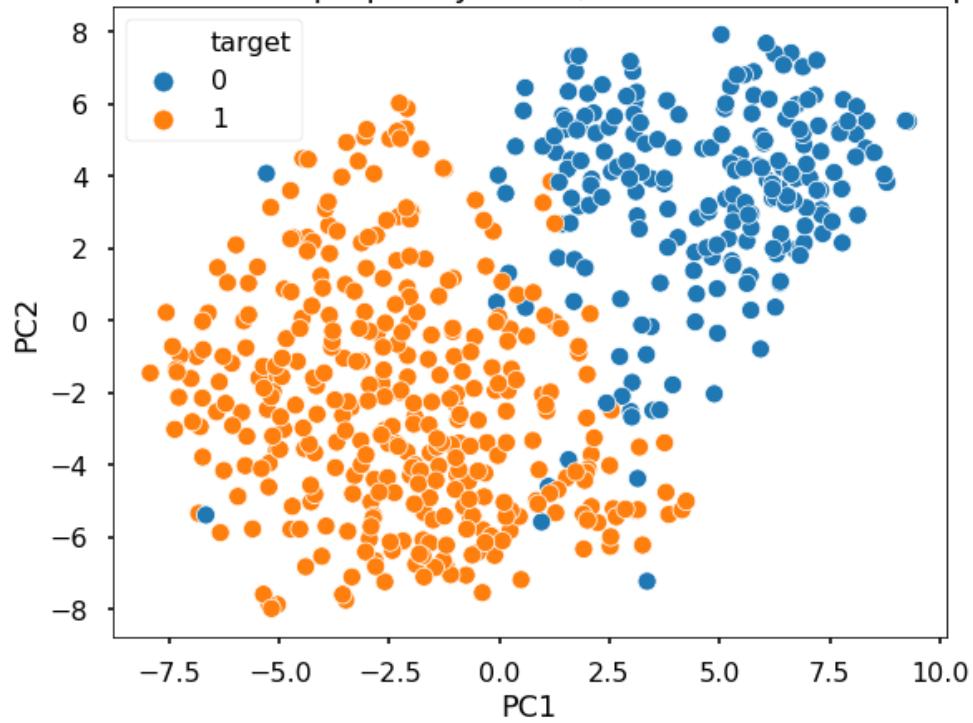
t-SNE visualization with perplexity -- 100, iterations -- 250 and epsilon -- 10



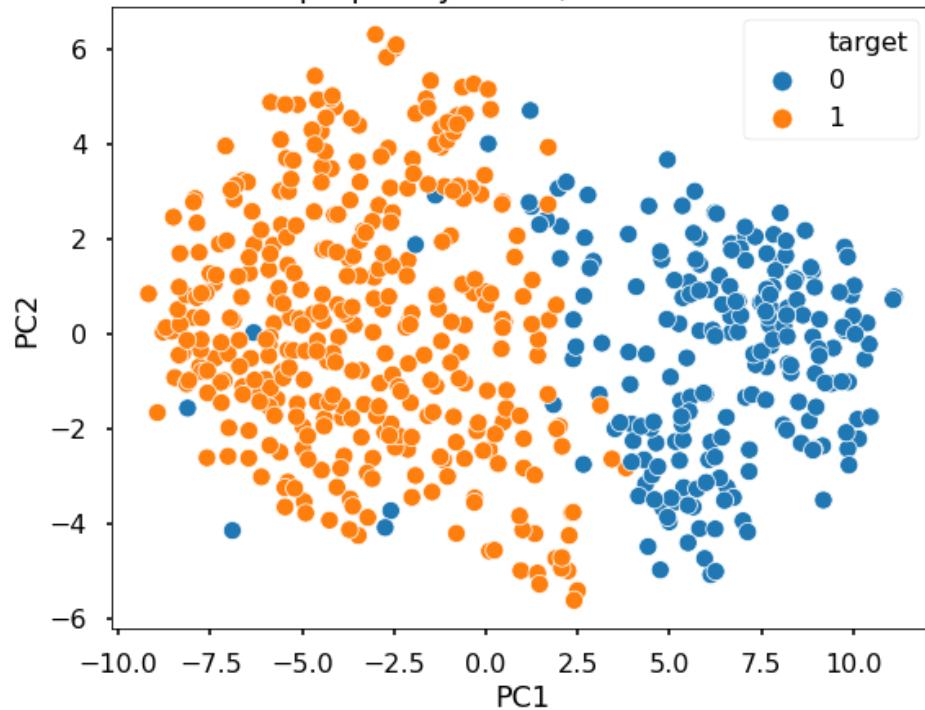
t-SNE visualization with perplexity -- 100, iterations -- 500 and epsilon -- 10



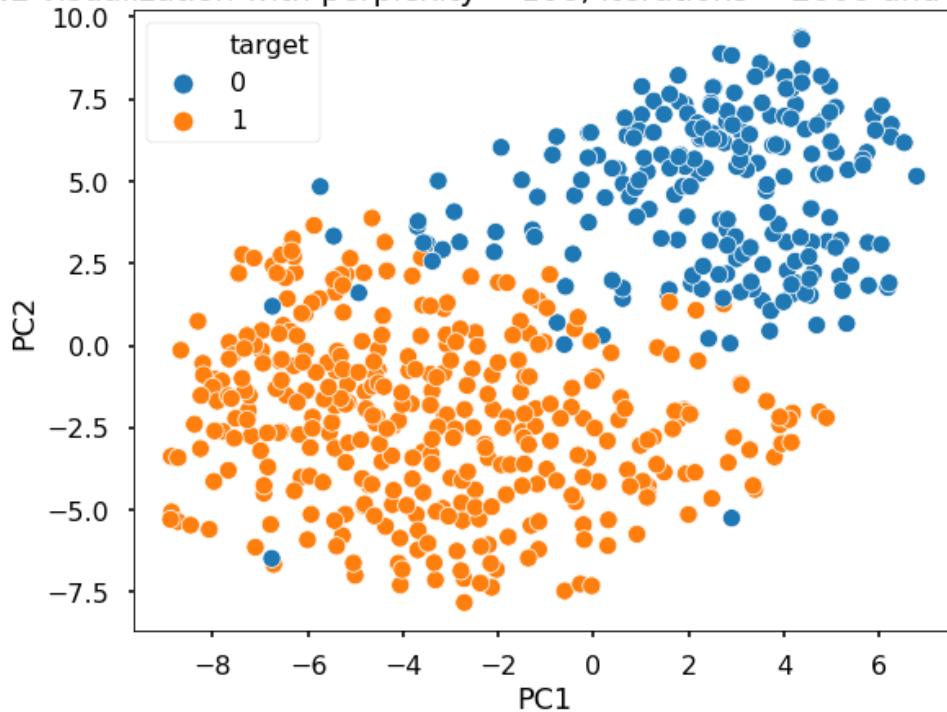
t-SNE visualization with perplexity -- 100, iterations -- 750 and epsilon -- 10



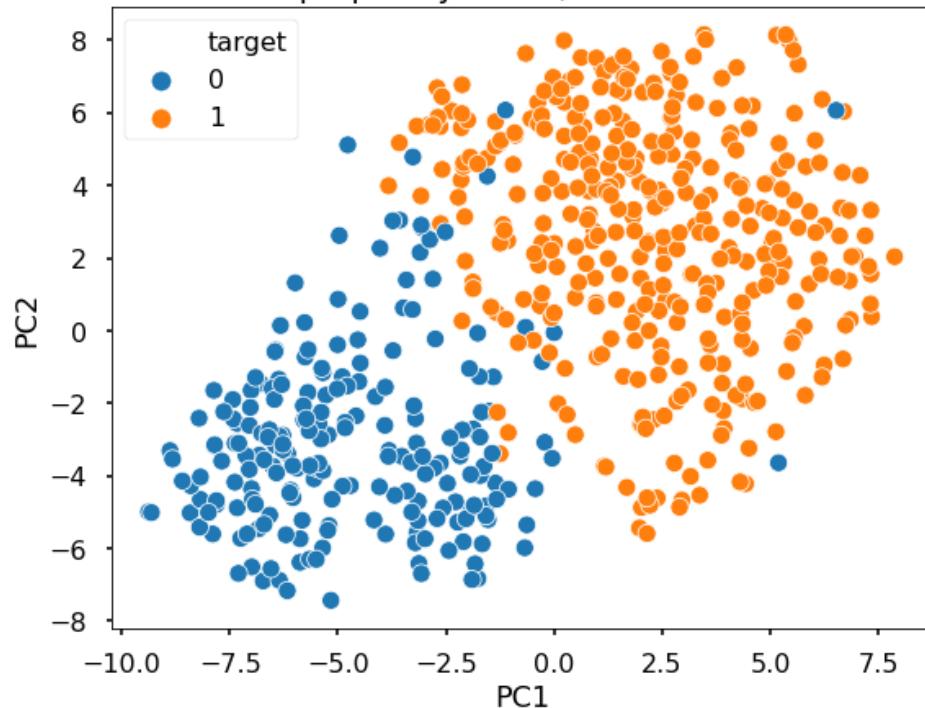
t-SNE visualization with perplexity -- 100, iterations -- 1000 and epsilon -- 10



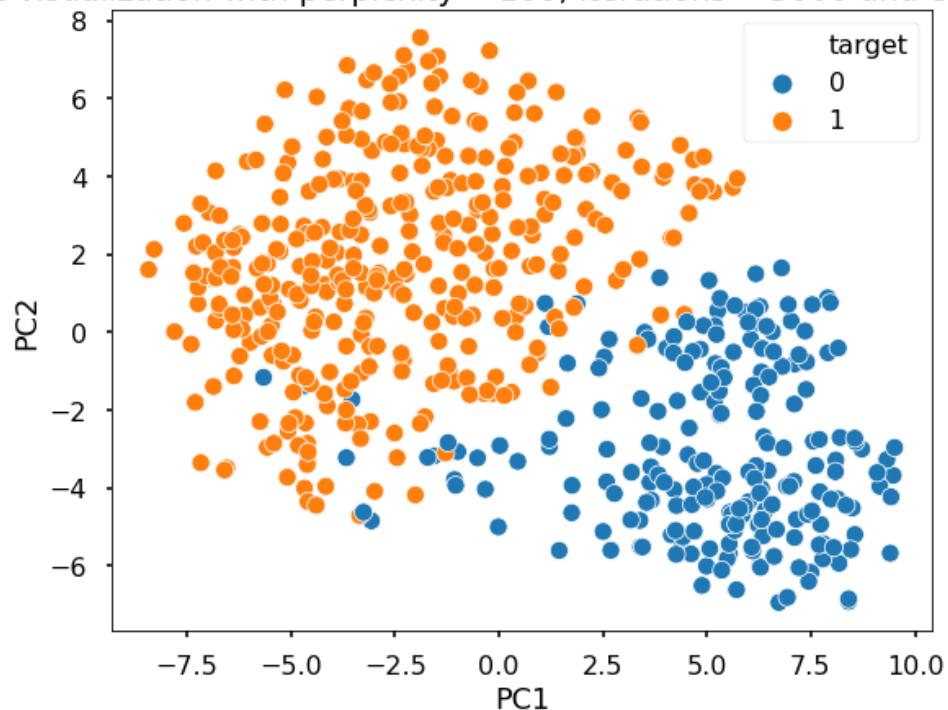
t-SNE visualization with perplexity -- 100, iterations -- 2000 and epsilon -- 10



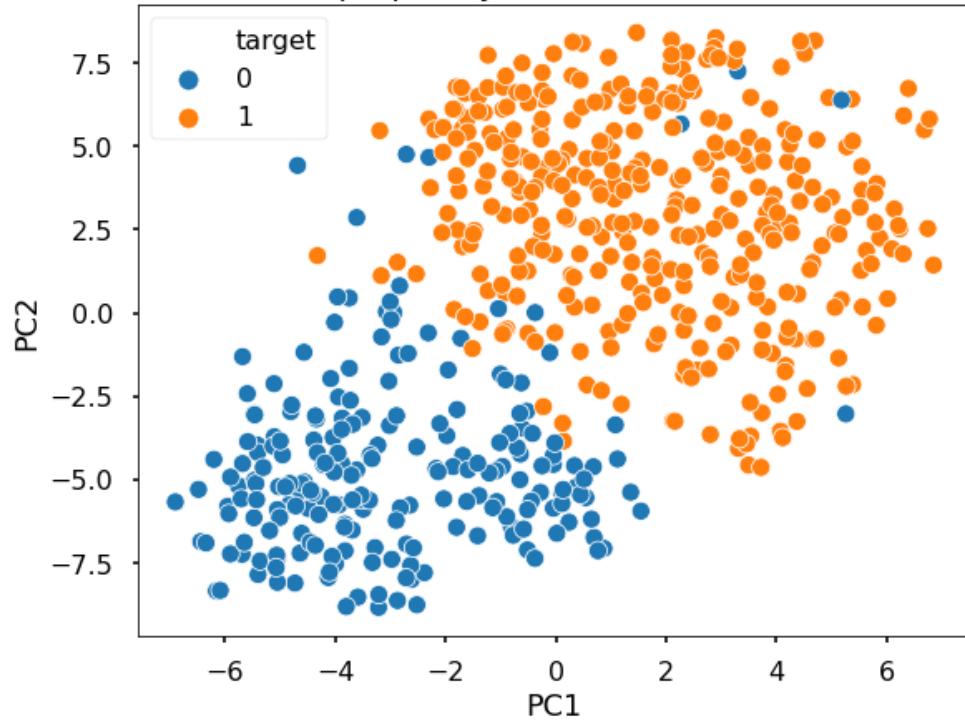
t-SNE visualization with perplexity -- 100, iterations -- 3000 and epsilon -- 10



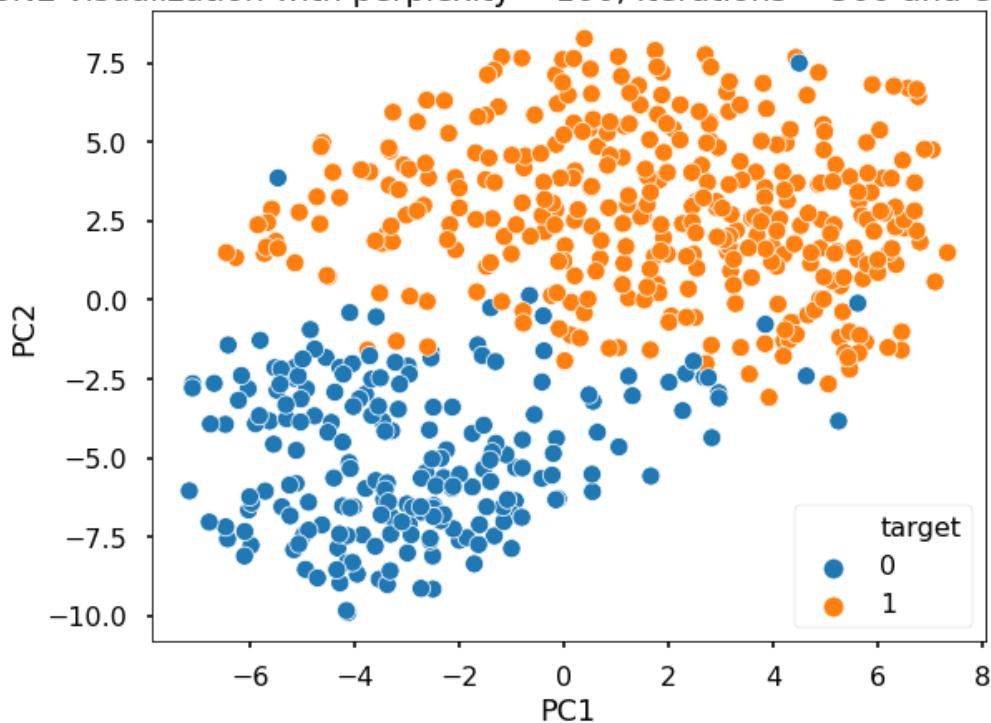
t-SNE visualization with perplexity -- 100, iterations -- 5000 and epsilon -- 10



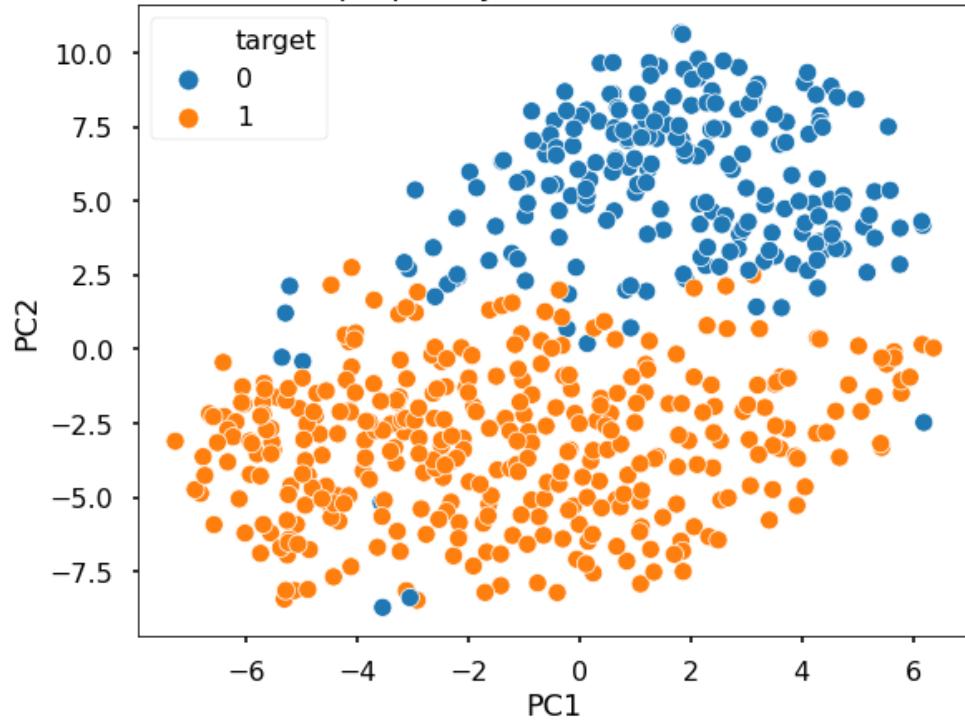
t-SNE visualization with perplexity -- 100, iterations -- 250 and epsilon -- 30



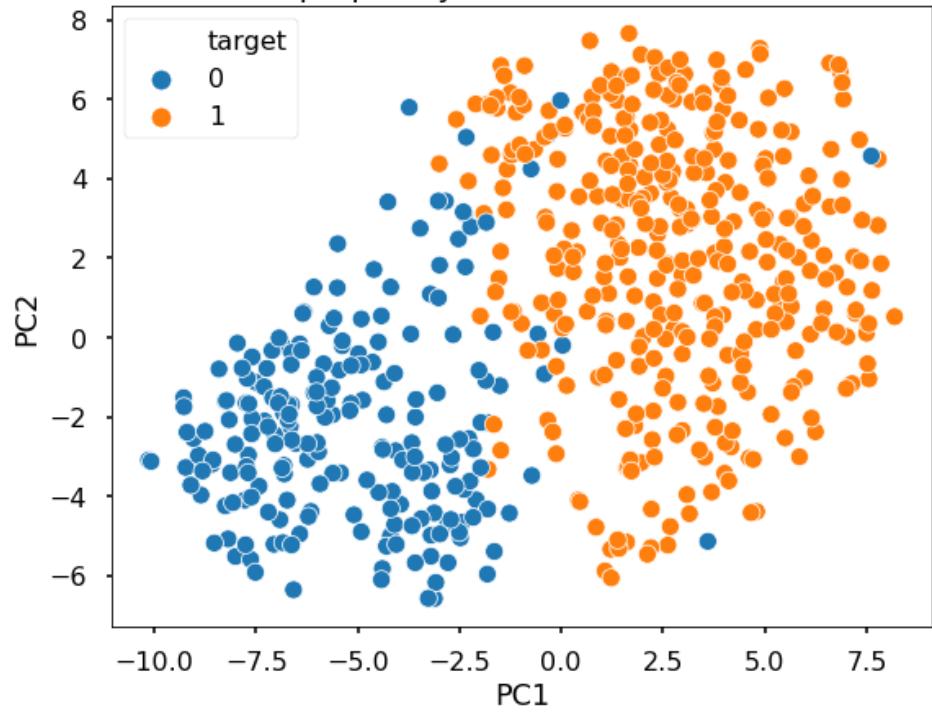
t-SNE visualization with perplexity -- 100, iterations -- 500 and epsilon -- 30



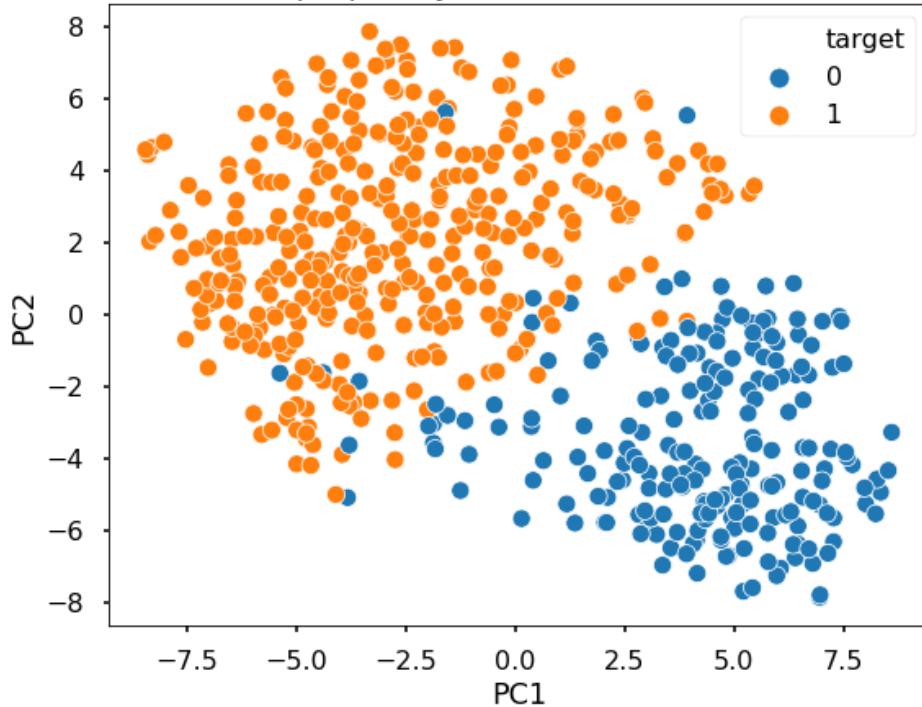
t-SNE visualization with perplexity -- 100, iterations -- 750 and epsilon -- 30



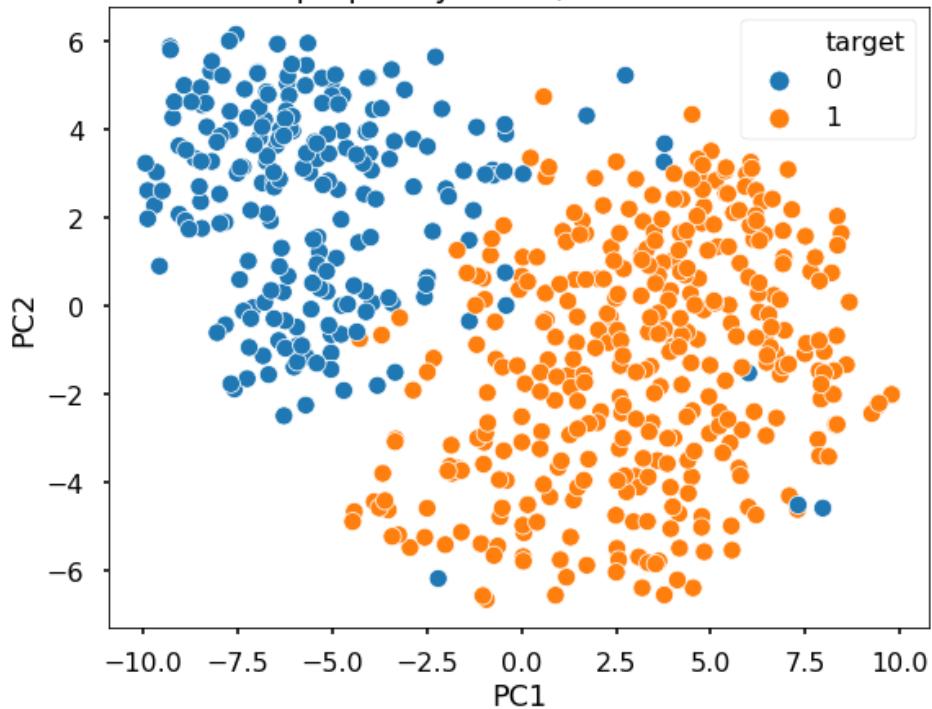
t-SNE visualization with perplexity -- 100, iterations -- 1000 and epsilon -- 30



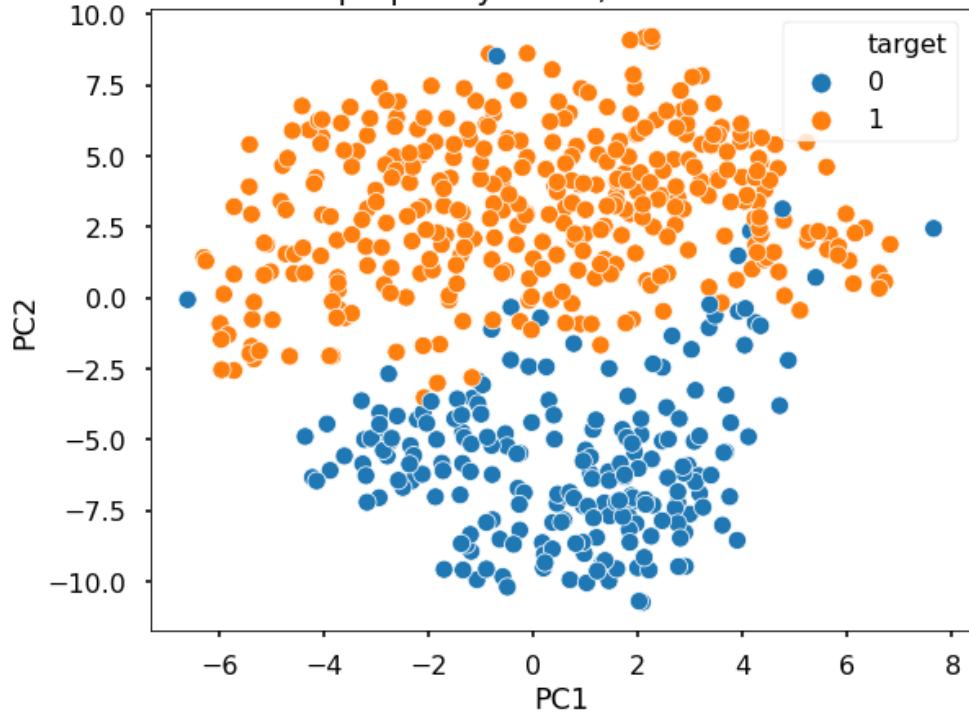
t-SNE visualization with perplexity -- 100, iterations -- 2000 and epsilon -- 30



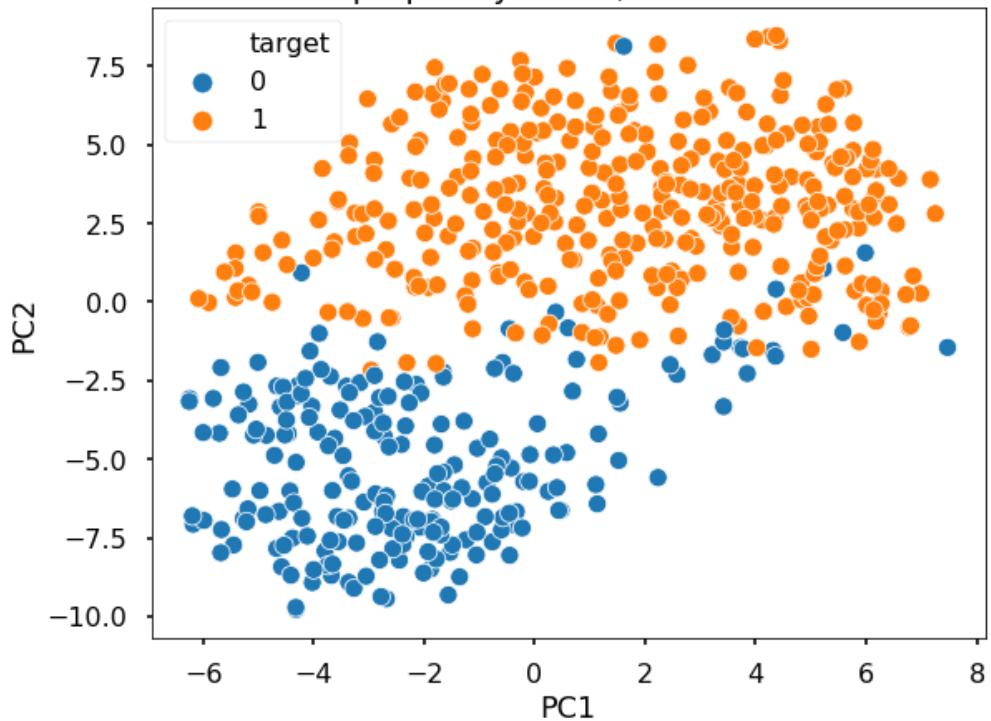
t-SNE visualization with perplexity -- 100, iterations -- 3000 and epsilon -- 30



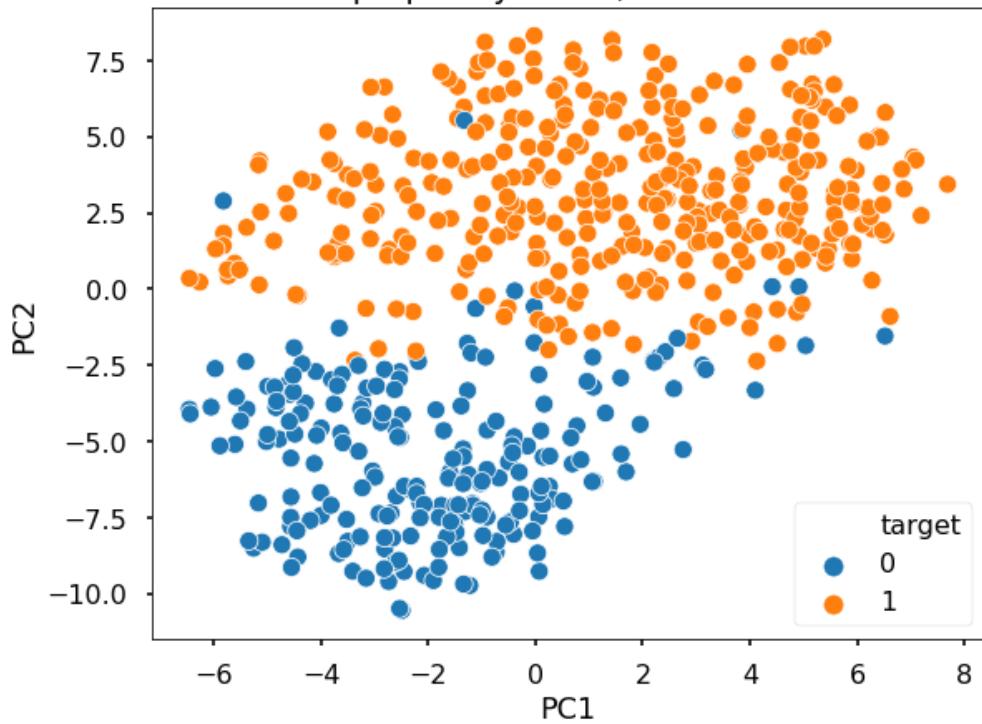
t-SNE visualization with perplexity -- 100, iterations -- 5000 and epsilon -- 30



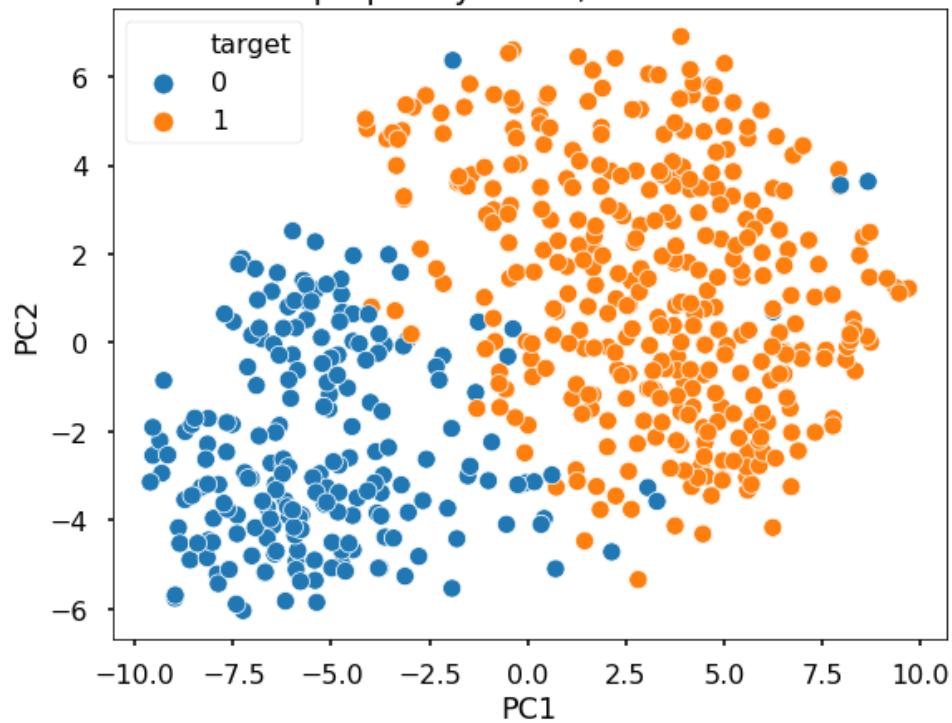
t-SNE visualization with perplexity -- 100, iterations -- 250 and epsilon -- 50



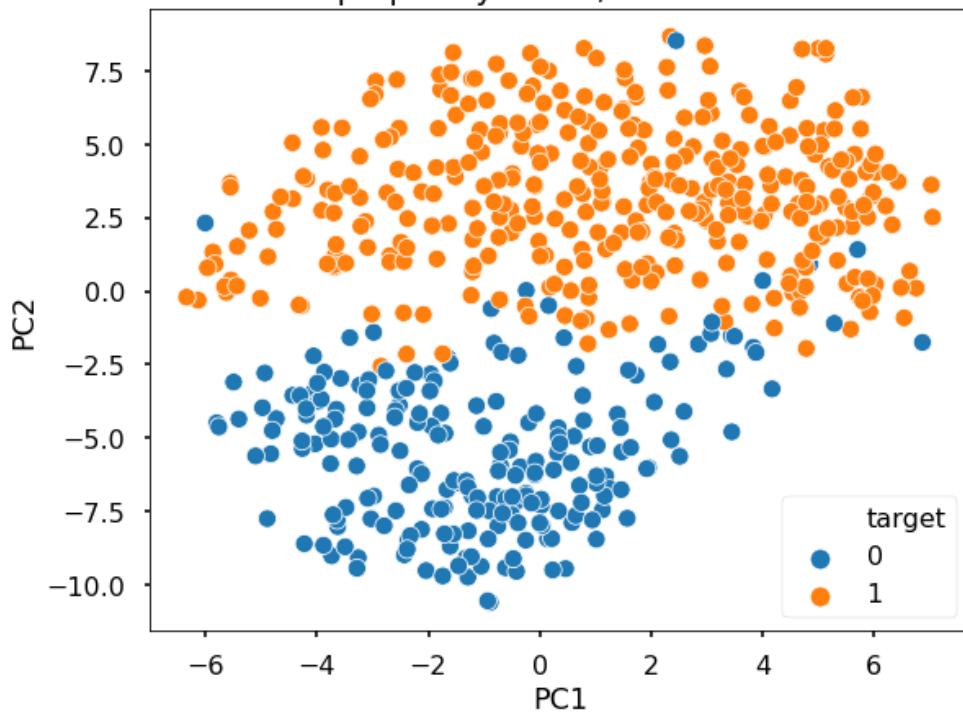
t-SNE visualization with perplexity -- 100, iterations -- 500 and epsilon -- 50



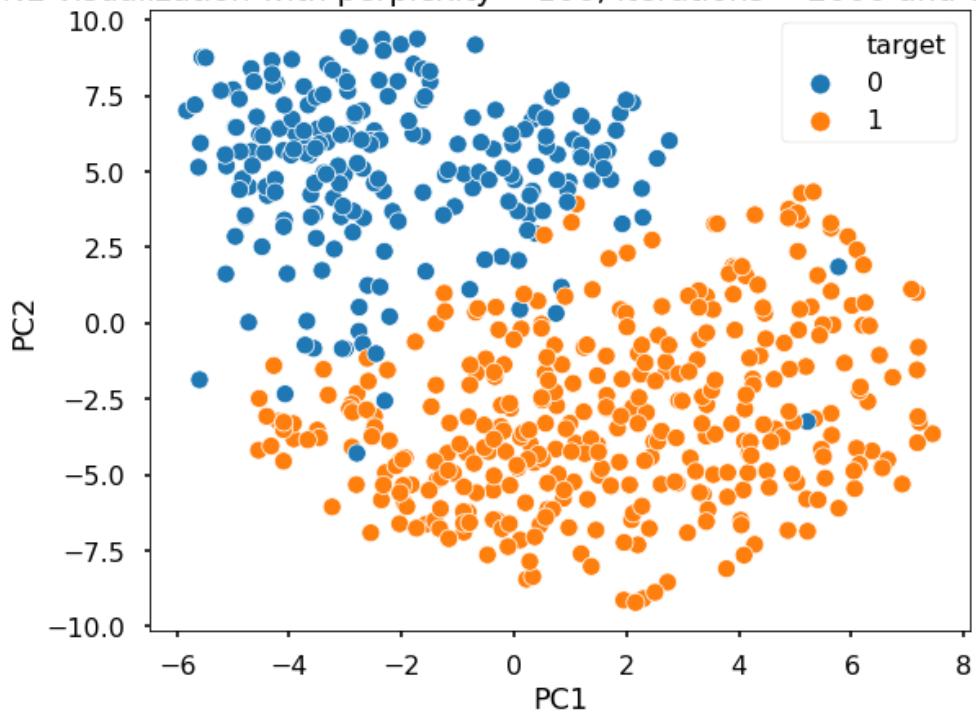
t-SNE visualization with perplexity -- 100, iterations -- 750 and epsilon -- 50



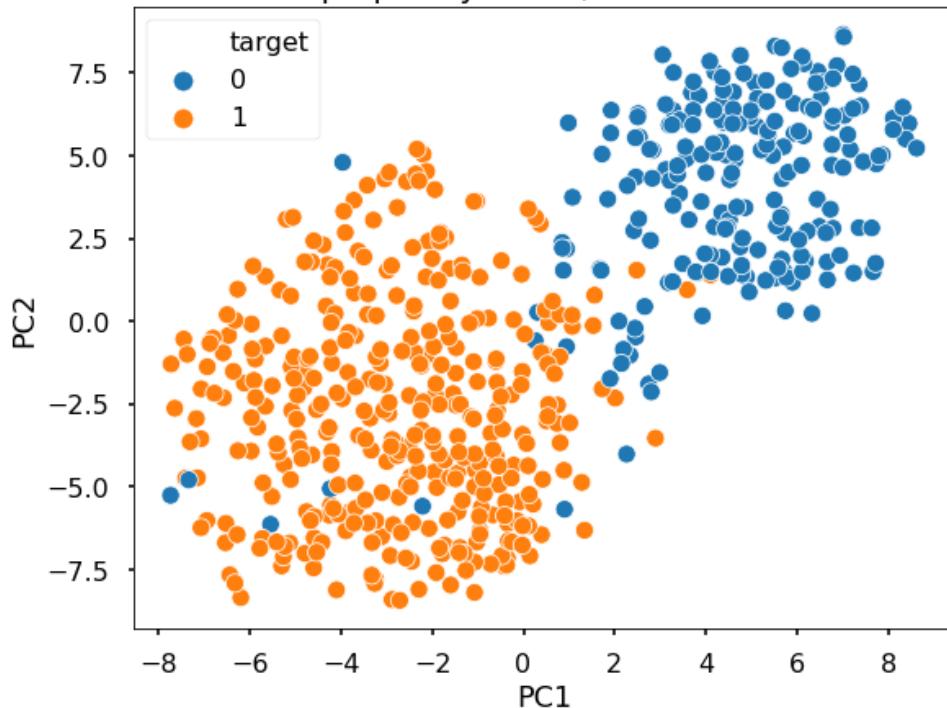
t-SNE visualization with perplexity -- 100, iterations -- 1000 and epsilon -- 50



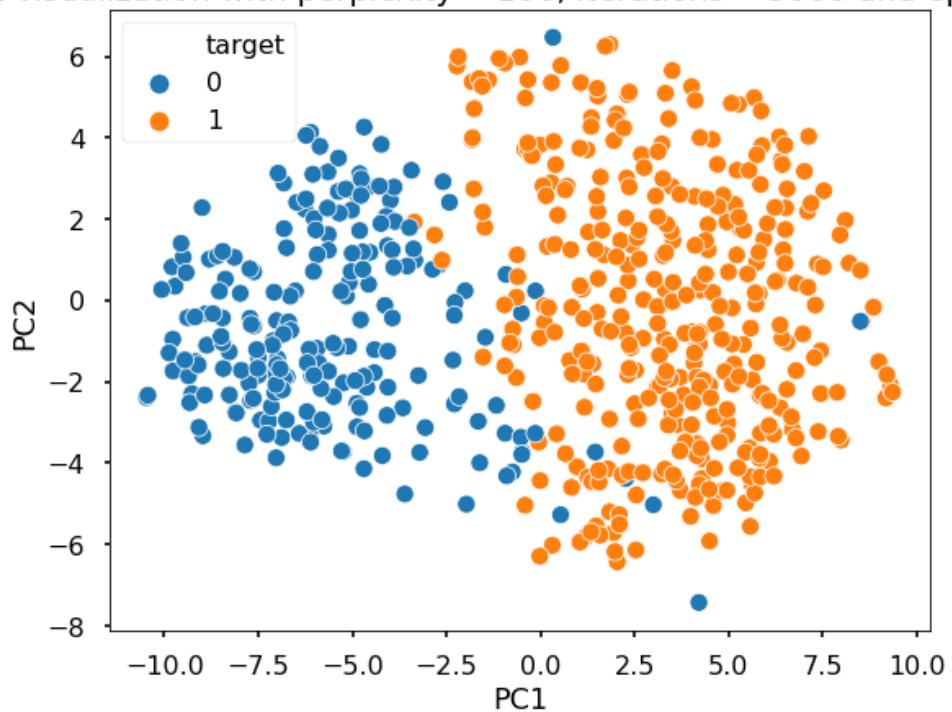
t-SNE visualization with perplexity -- 100, iterations -- 2000 and epsilon -- 50



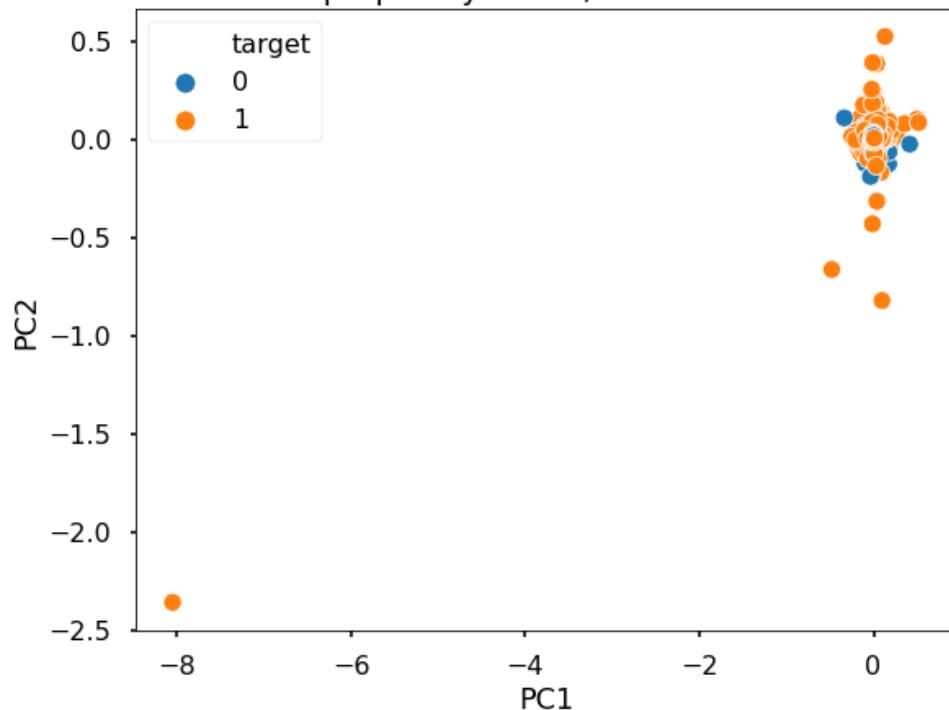
t-SNE visualization with perplexity -- 100, iterations -- 3000 and epsilon -- 50



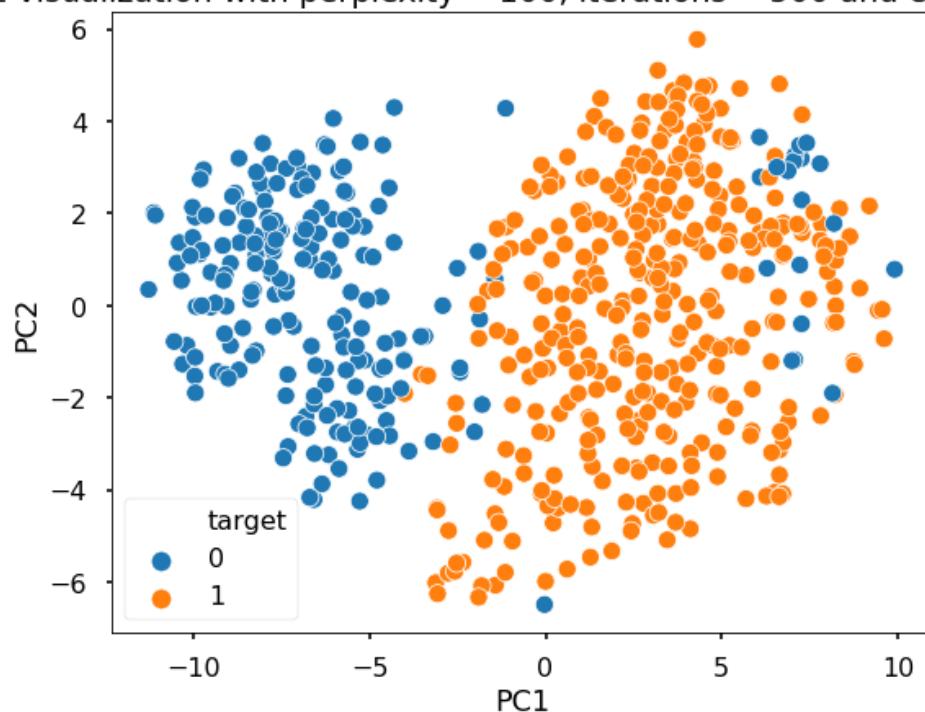
t-SNE visualization with perplexity -- 100, iterations -- 5000 and epsilon -- 50



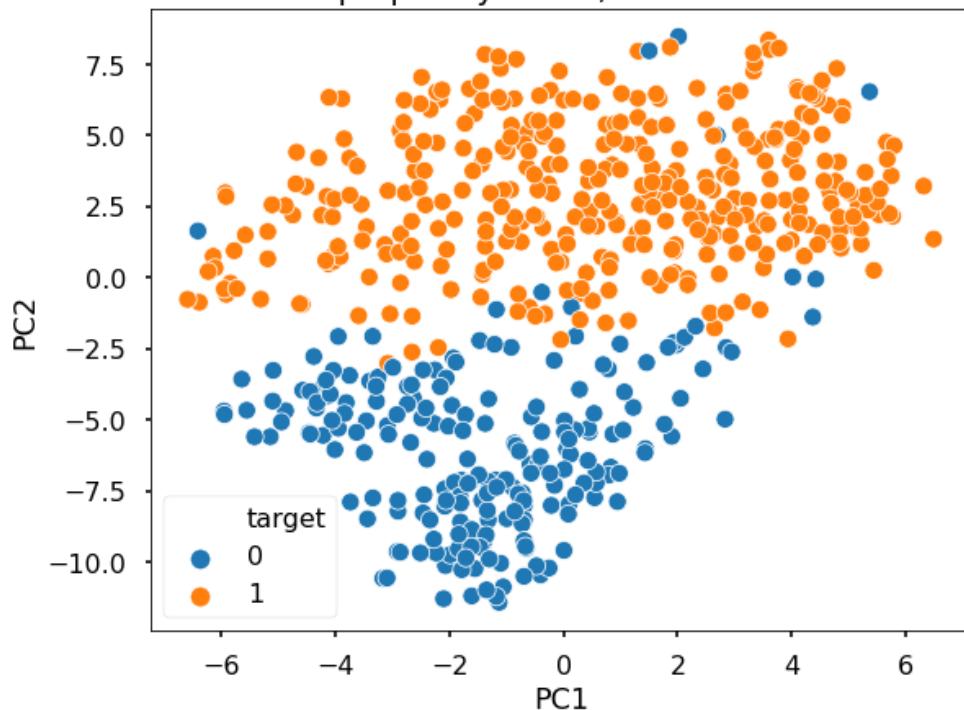
t-SNE visualization with perplexity -- 100, iterations -- 250 and epsilon -- 100



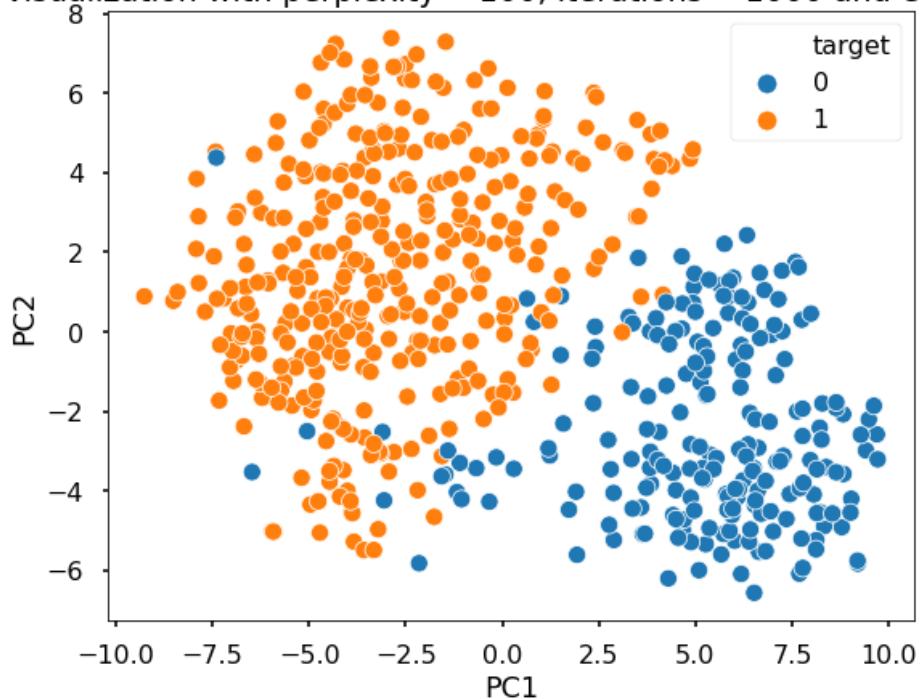
t-SNE visualization with perplexity -- 100, iterations -- 500 and epsilon -- 100



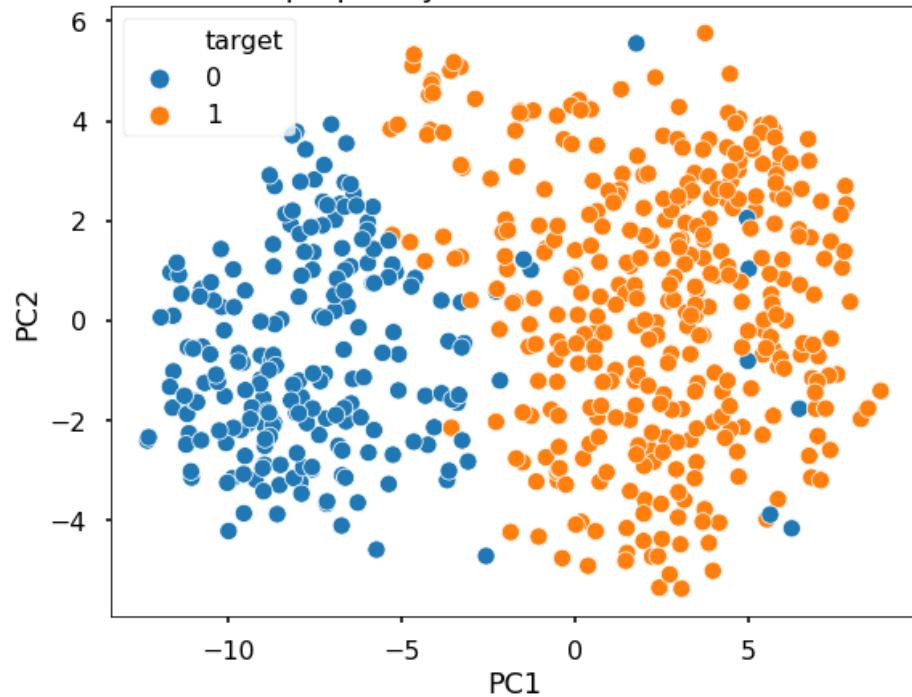
t-SNE visualization with perplexity -- 100, iterations -- 750 and epsilon -- 100



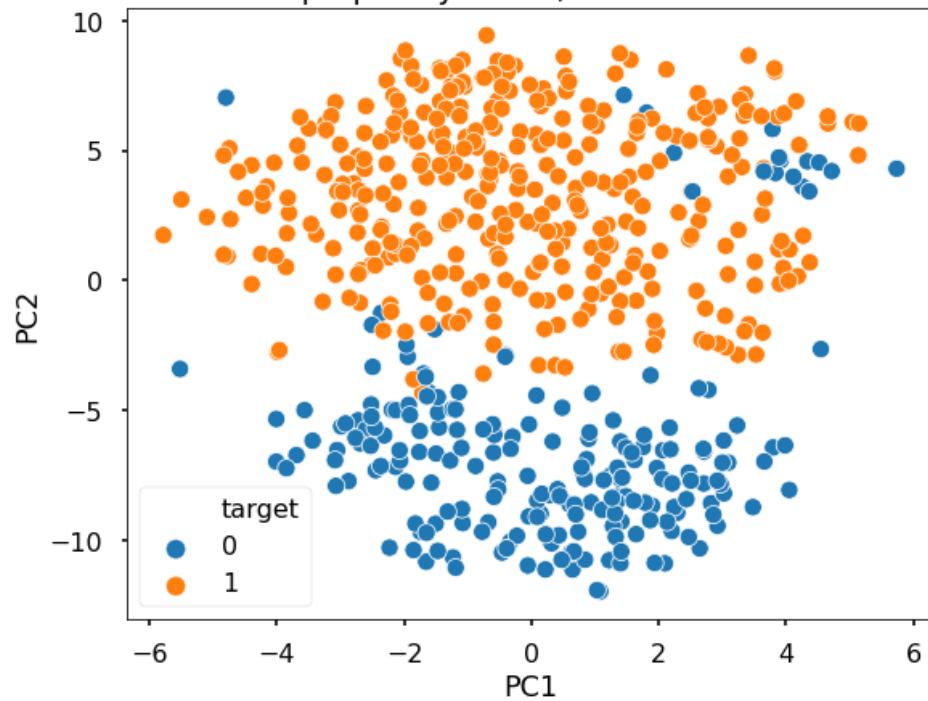
t-SNE visualization with perplexity -- 100, iterations -- 1000 and epsilon -- 100



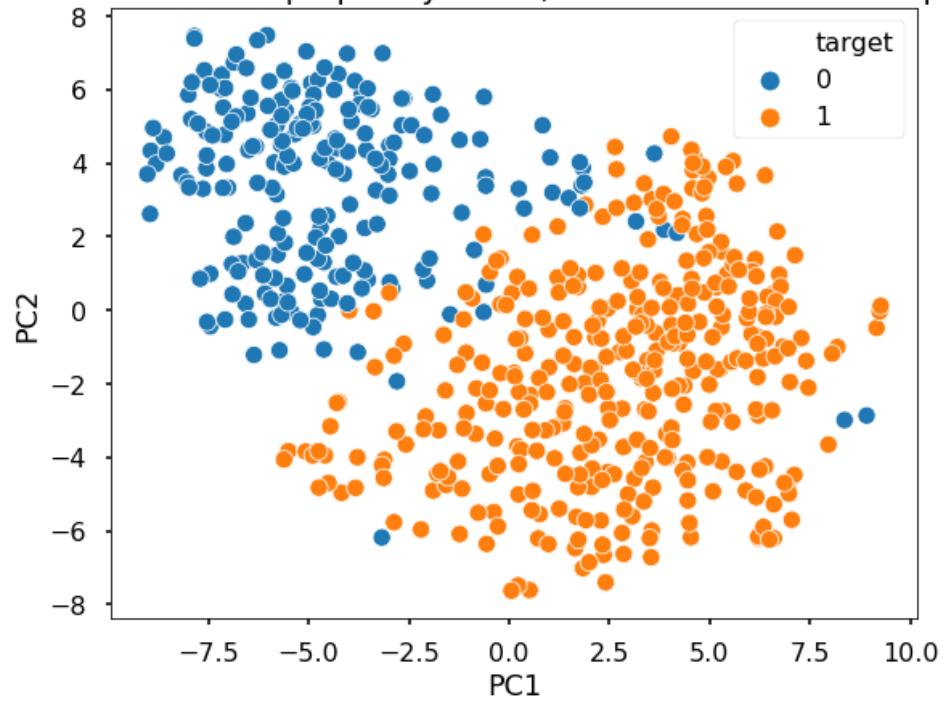
t-SNE visualization with perplexity -- 100, iterations -- 2000 and epsilon -- 100



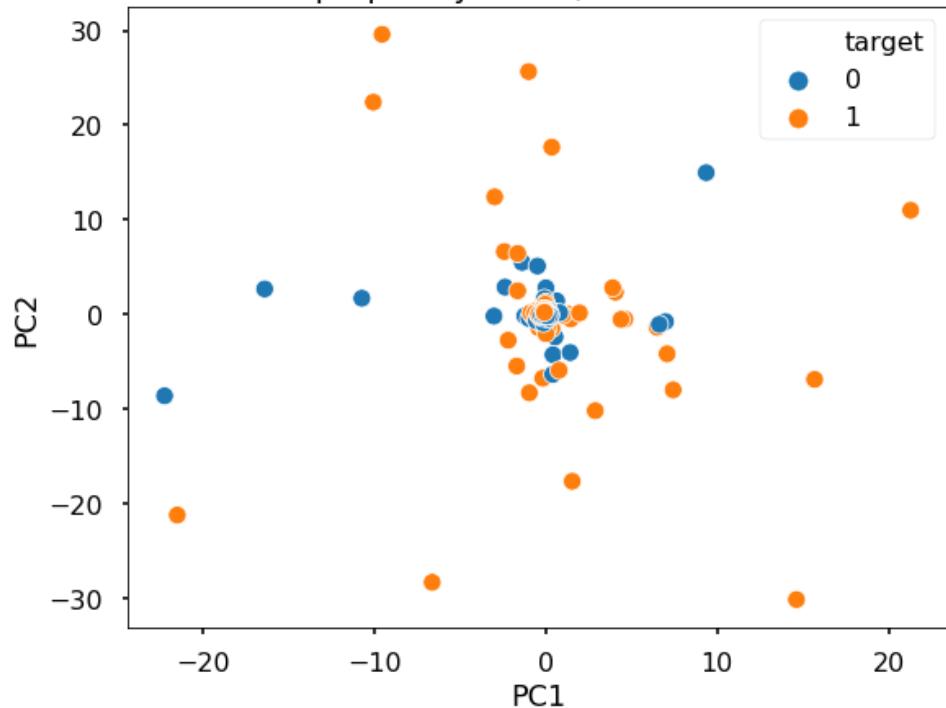
t-SNE visualization with perplexity -- 100, iterations -- 3000 and epsilon -- 100



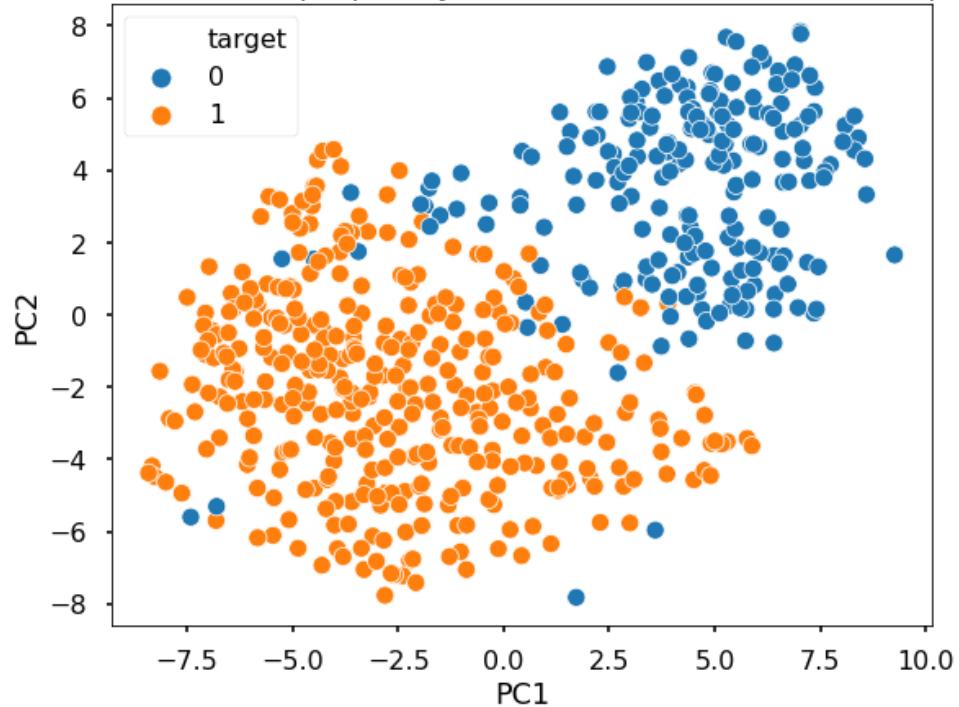
t-SNE visualization with perplexity -- 100, iterations -- 5000 and epsilon -- 100



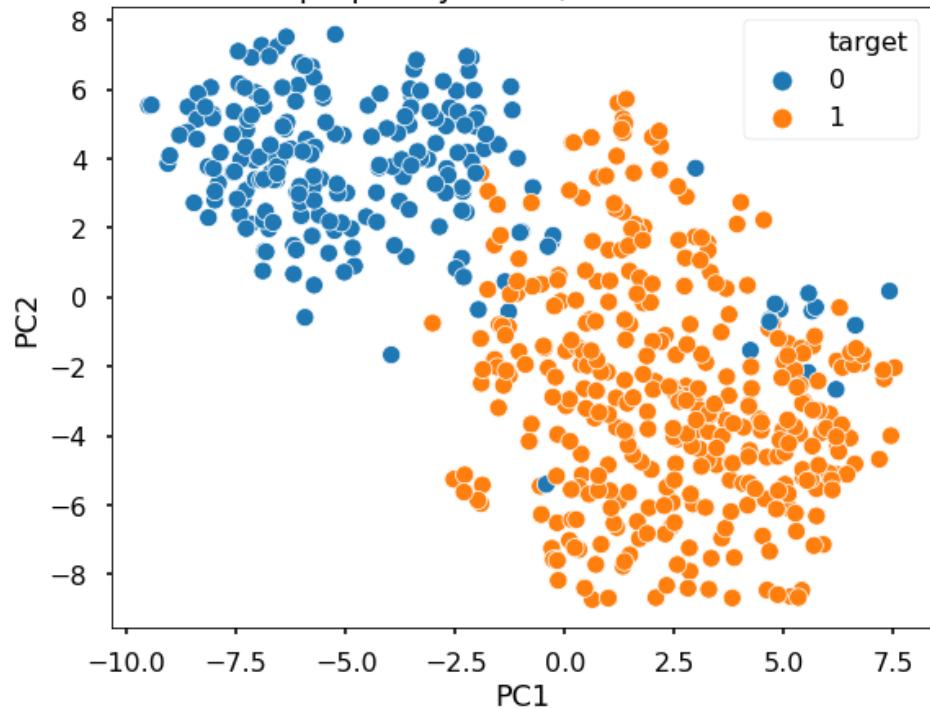
t-SNE visualization with perplexity -- 100, iterations -- 250 and epsilon -- 200



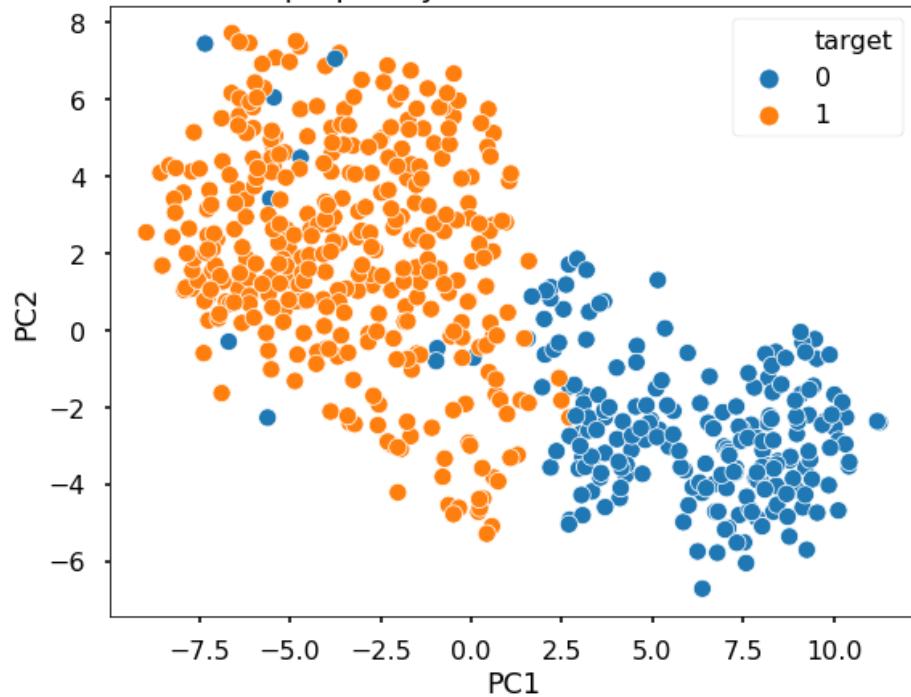
t-SNE visualization with perplexity -- 100, iterations -- 500 and epsilon -- 200



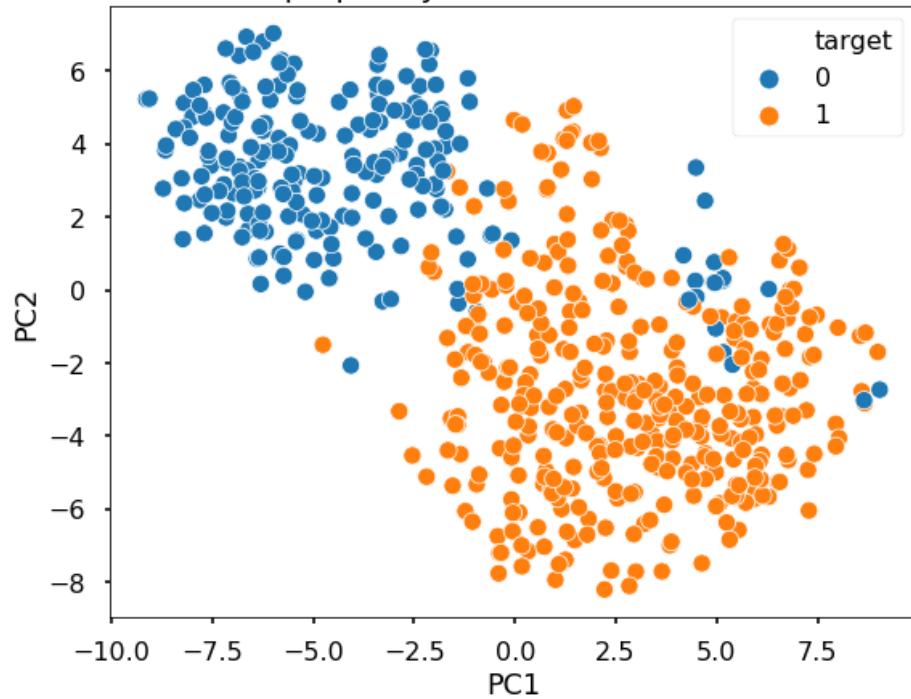
t-SNE visualization with perplexity -- 100, iterations -- 750 and epsilon -- 200



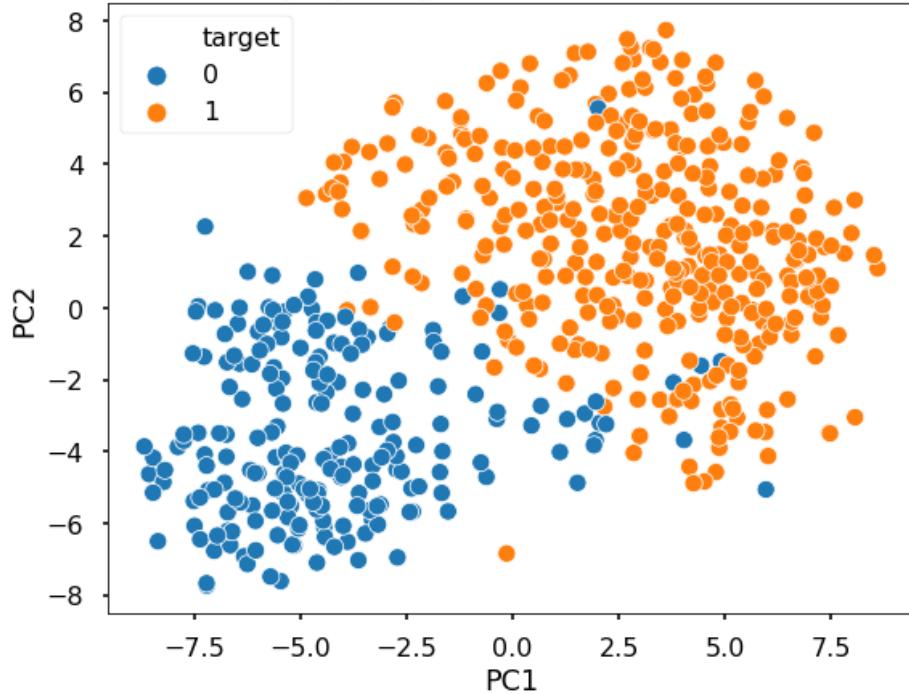
t-SNE visualization with perplexity -- 100, iterations -- 1000 and epsilon -- 200



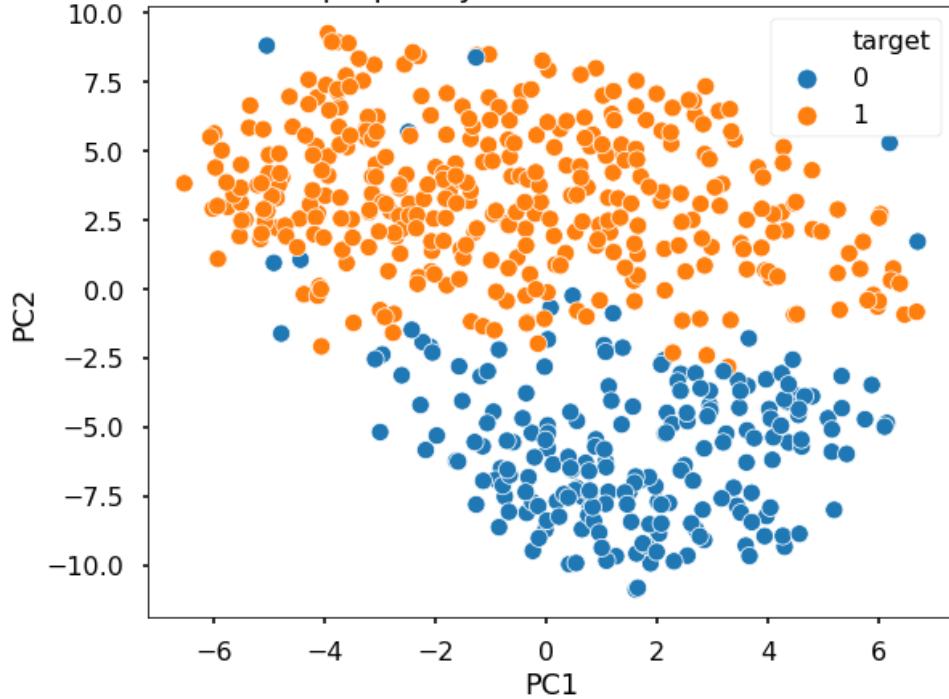
t-SNE visualization with perplexity -- 100, iterations -- 2000 and epsilon -- 200



t-SNE visualization with perplexity -- 100, iterations -- 3000 and epsilon -- 200



t-SNE visualization with perplexity -- 100, iterations -- 5000 and epsilon -- 200



In the above plots, my observation is that with Perplexity=30 and Epsilon=30 gives us the better separation of classes with a very few outliers. However, in the other plots the outliers are quite high and they are largely distant from the respective class cluster.

So, next step is to run the T-SNE for this combination multiple times with various iterations.

CASE-II-Multiple_runs

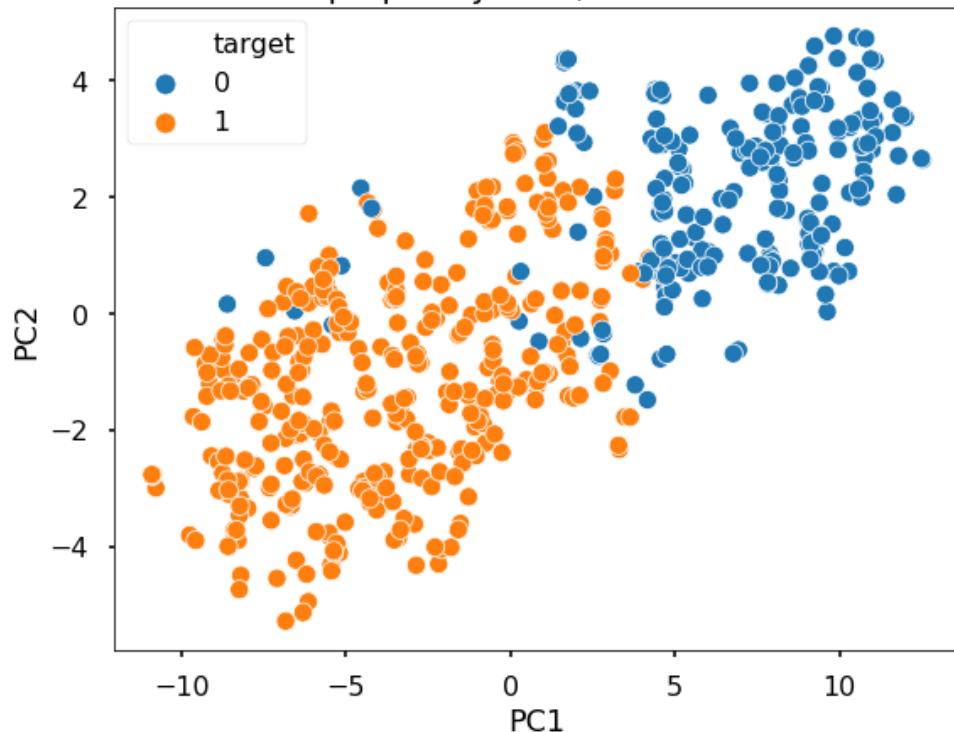
Running t-SNE for a range of iterations with Learning rate and perplexity as 30 and embedding method as 'random'

```
In [7]: iterations = [250,500,750,1000,2000,3500,5000]

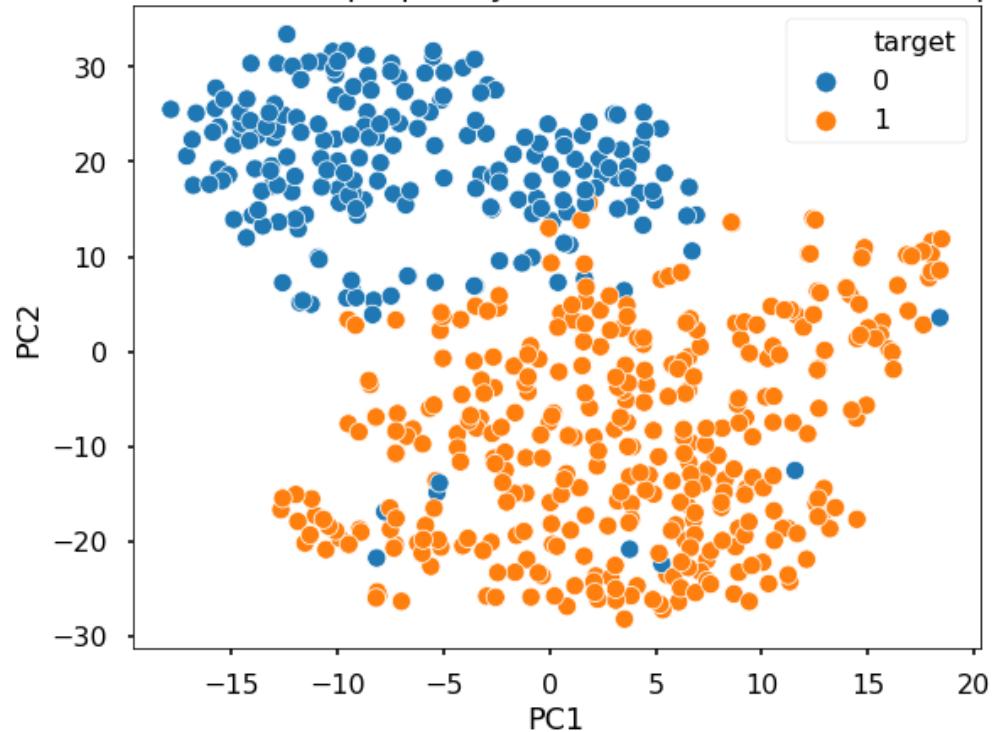
for idx in range(len(iterations)):
    tsne1 = TSNE(n_components=2,init='random',perplexity=30,learning_rate=30,n_iter=
                  method='exact',
                  n_jobs=-1)
    cancer_tsne_pcmps = pd.DataFrame(tsne1.fit_transform(cancer_norm_df),columns=[ 'P
    cancer_tsne_pcmps = pd.concat([cancer_tsne_pcmps,cancer_df[ 'target']],axis=1)
    with plt.style.context('seaborn-poster'):
        plt.figure(figsize=(9,7))
        sns.scatterplot(data=cancer_tsne_pcmps,x='PC1',y='PC2',hue='target')
        plt.title("t-SNE visualization with perplexity -- {0}, iterations -- {1} and

plt.show()
```

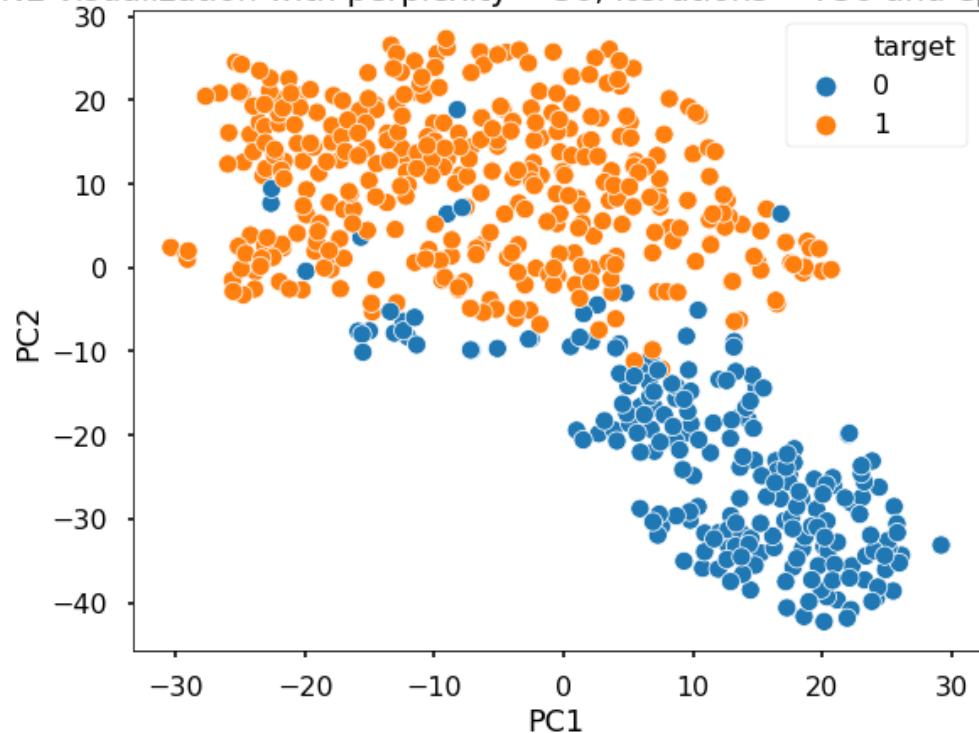
t-SNE visualization with perplexity -- 30, iterations -- 250 and epsilon -- 30



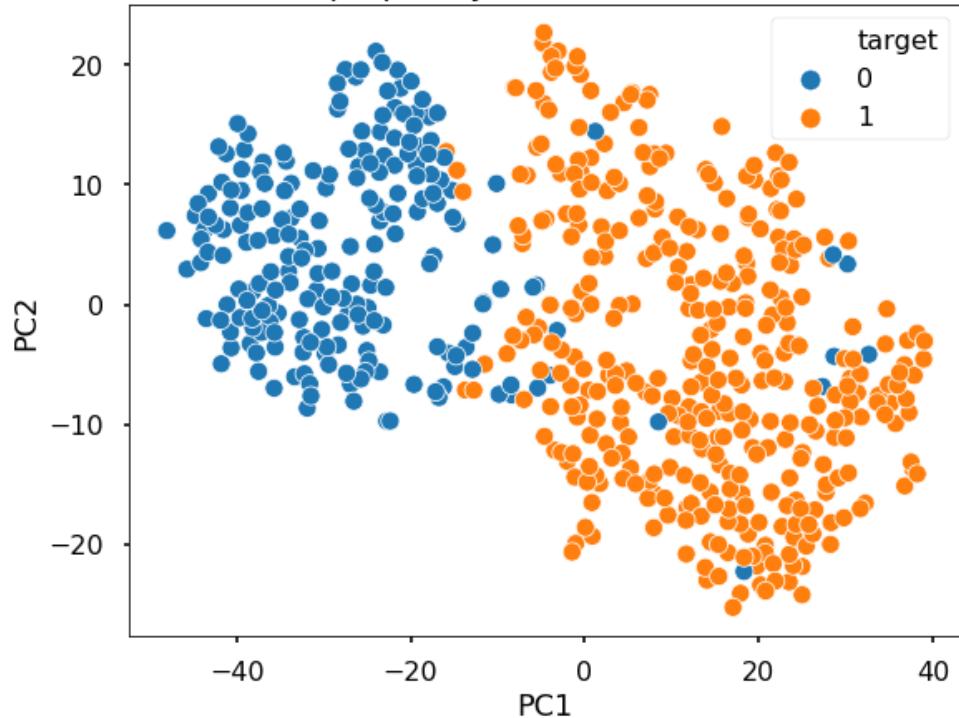
t-SNE visualization with perplexity -- 30, iterations -- 500 and epsilon -- 30



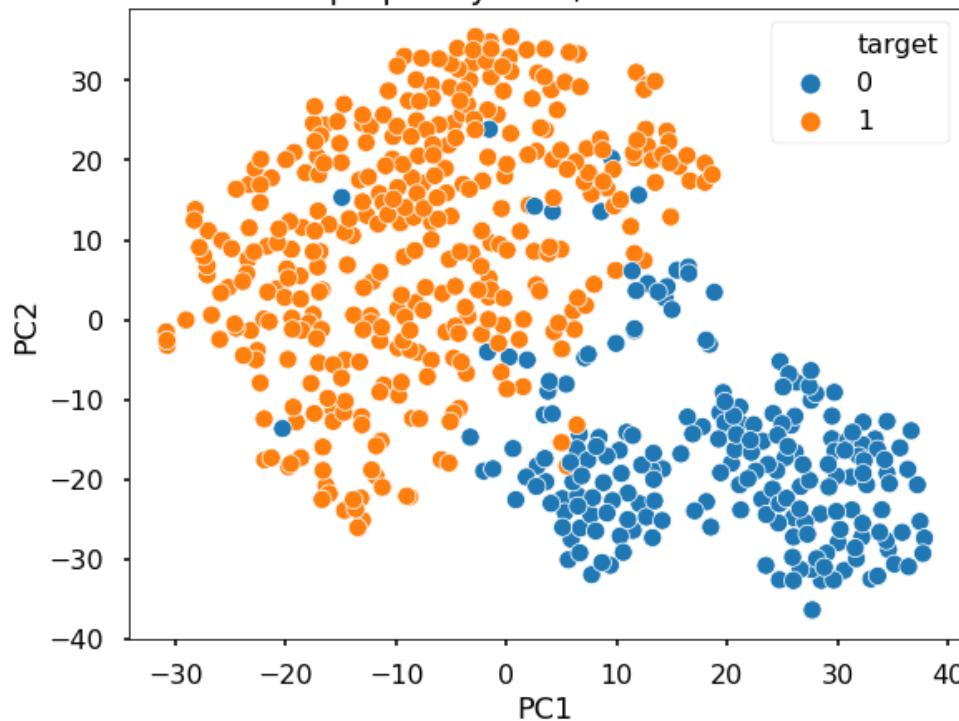
t-SNE visualization with perplexity -- 30, iterations -- 750 and epsilon -- 30



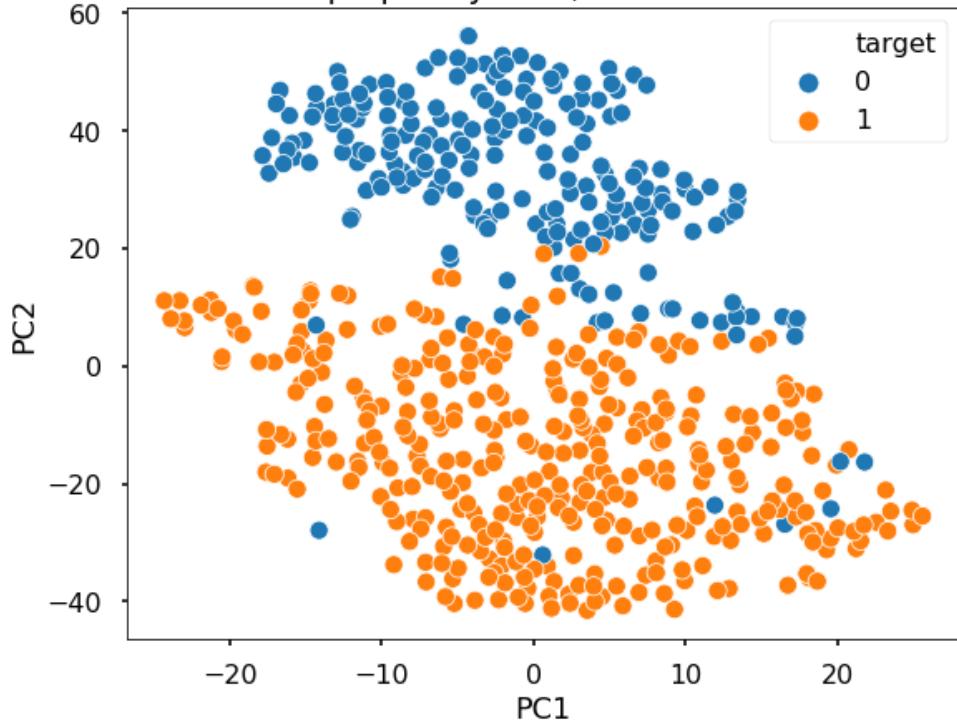
t-SNE visualization with perplexity -- 30, iterations -- 1000 and epsilon -- 30



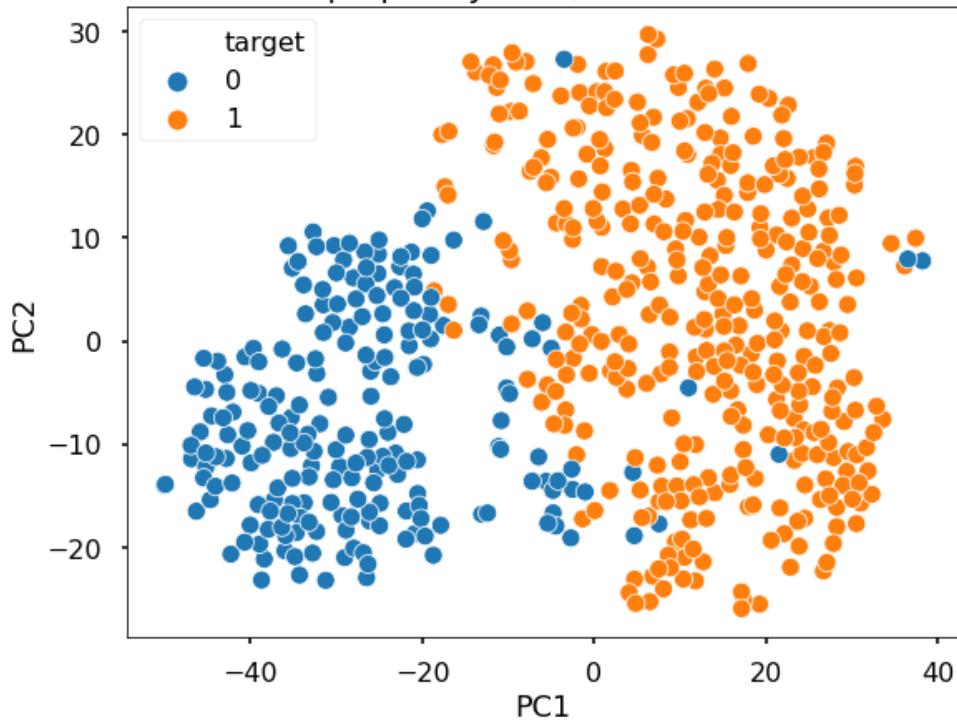
t-SNE visualization with perplexity -- 30, iterations -- 2000 and epsilon -- 30



t-SNE visualization with perplexity -- 30, iterations -- 3500 and epsilon -- 30



t-SNE visualization with perplexity -- 30, iterations -- 5000 and epsilon -- 30



CASE-III-Multiple_runs

Running t-SNE for a range of iterations with Learning rate and perplexity as 30 and embedding method as 'PCA'

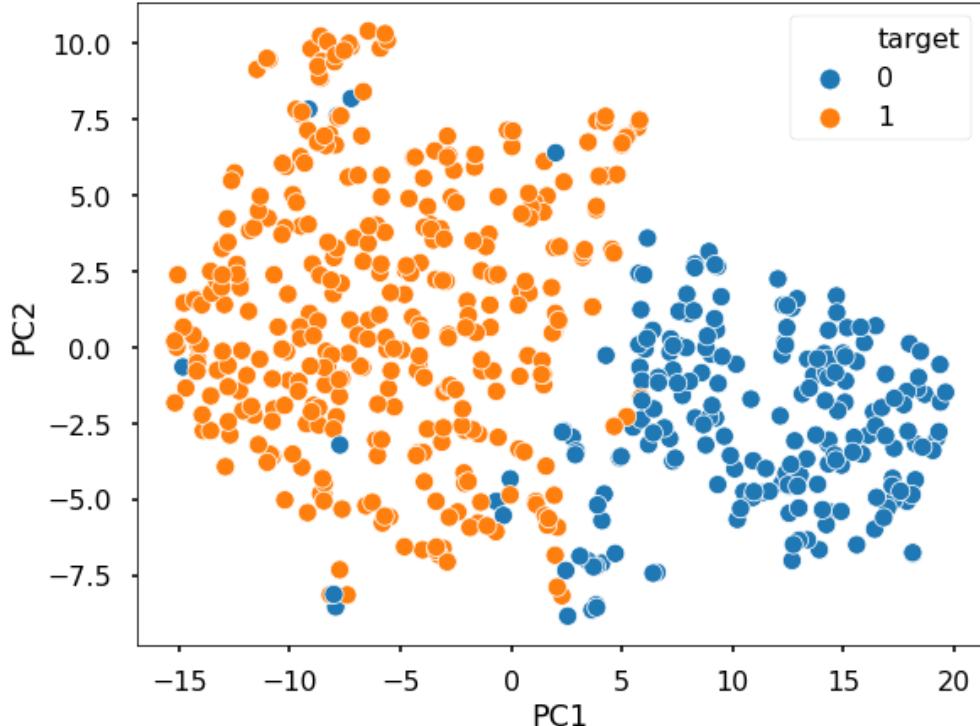
```
In [8]: iterations = [250,500,750,1000,2000,3500,5000]
for idx in range(len(iterations)):
    tsne1 = TSNE(n_components=2,init='pca',perplexity=30,learning_rate=30,n_iter=ite
               method='exact',
```

```

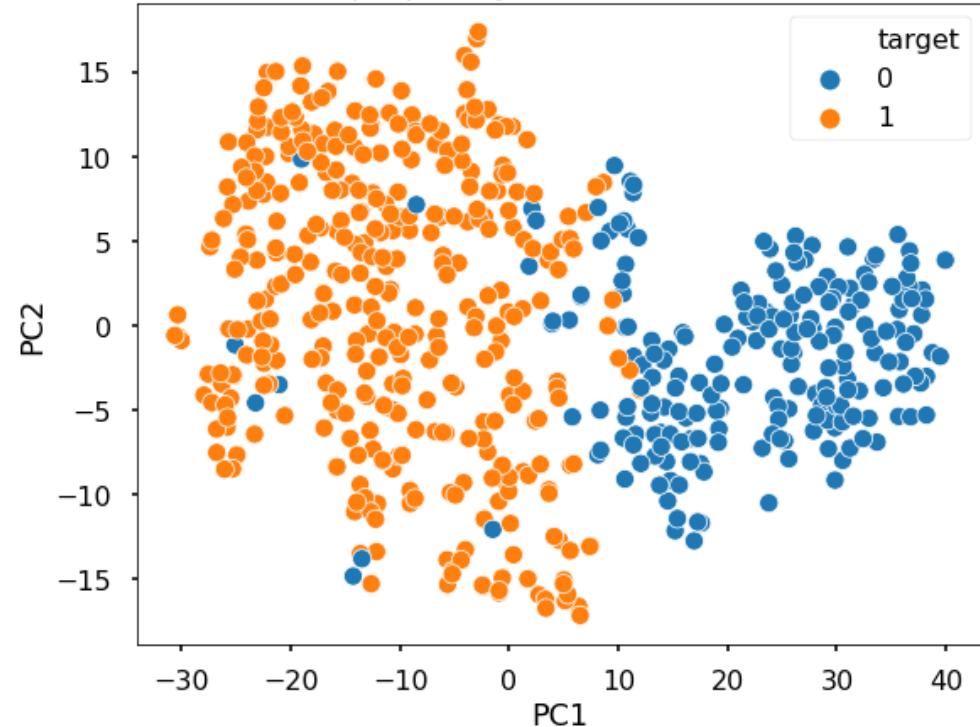
n_jobs=-1)
cancer_tsne_pcmps = pd.DataFrame(tsne1.fit_transform(cancer_norm_df),columns=['PC1','PC2'])
cancer_tsne_pcmps = pd.concat([cancer_tsne_pcmps,cancer_df['target']],axis=1)
with plt.style.context('seaborn-poster'):
    plt.figure(figsize=(9,7))
    sns.scatterplot(data=cancer_tsne_pcmps,x='PC1',y='PC2',hue='target')
    plt.title("t-SNE visualization with perplexity -- {0}, iterations -- {1} and epsilon -- {2}")
plt.show()

```

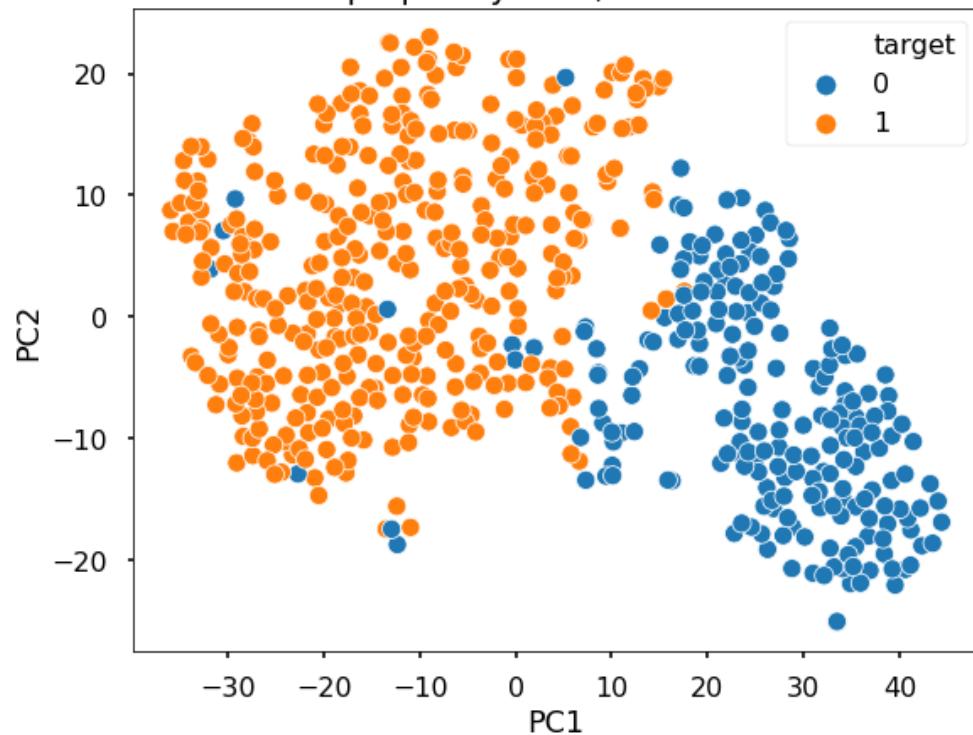
t-SNE visualization with perplexity -- 30, iterations -- 250 and epsilon -- 30



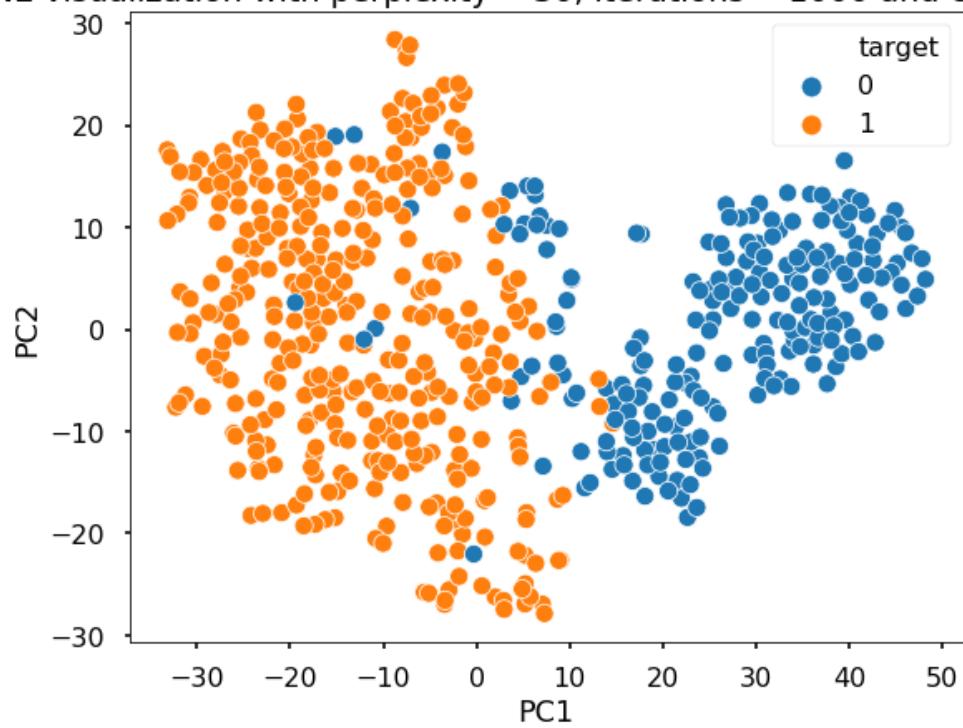
t-SNE visualization with perplexity -- 30, iterations -- 500 and epsilon -- 30



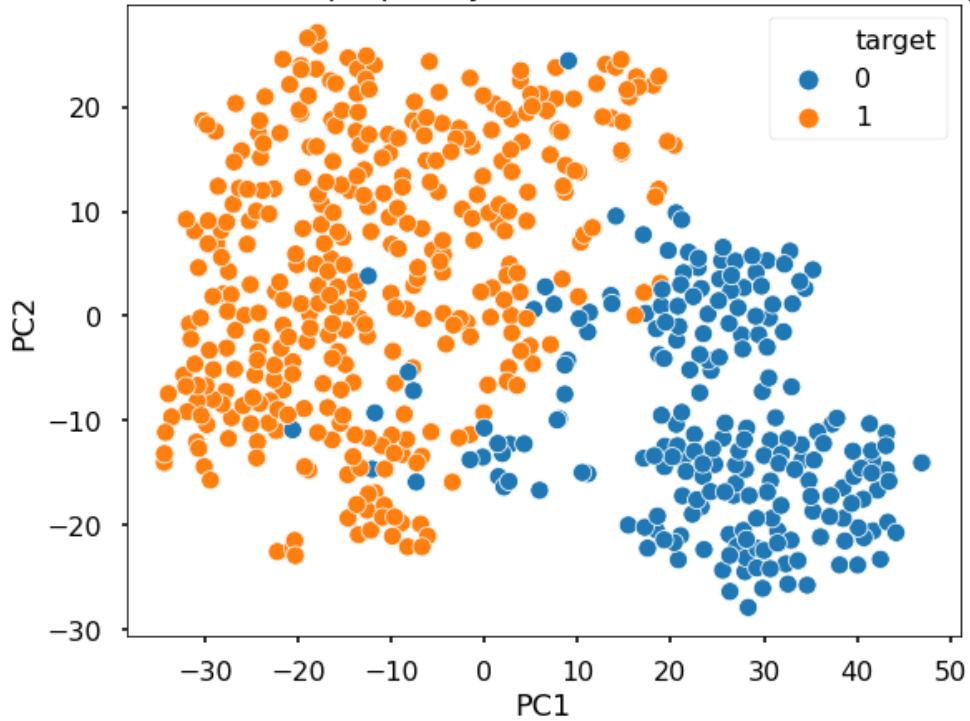
t-SNE visualization with perplexity -- 30, iterations -- 750 and epsilon -- 30



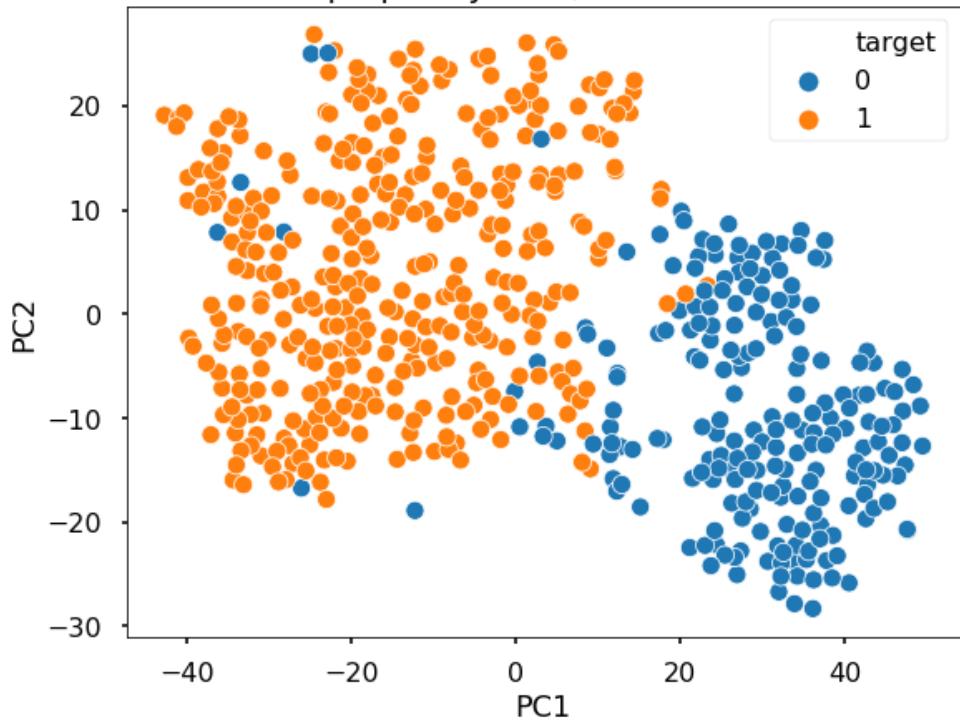
t-SNE visualization with perplexity -- 30, iterations -- 1000 and epsilon -- 30



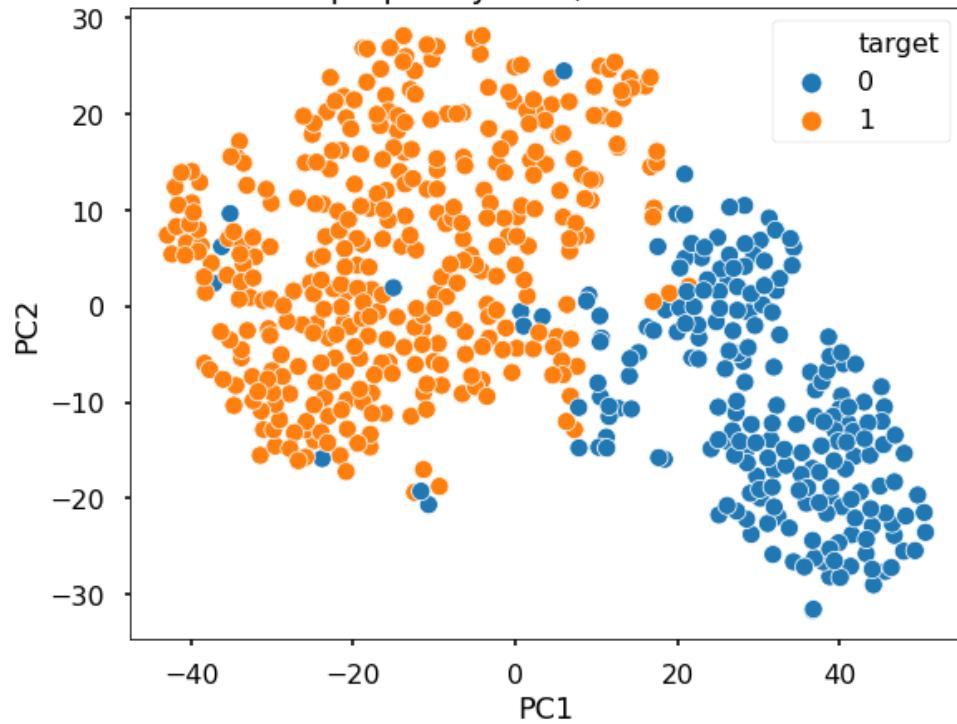
t-SNE visualization with perplexity -- 30, iterations -- 2000 and epsilon -- 30



t-SNE visualization with perplexity -- 30, iterations -- 3500 and epsilon -- 30



t-SNE visualization with perplexity -- 30, iterations -- 5000 and epsilon -- 30



IRIS_Dataset

```
In [26]: iris_df = pd.DataFrame(iris.data,columns=iris.feature_names)
iris_stand_df = pd.concat([pd.DataFrame(ss.fit_transform(iris_df),columns=iris.feature_names),
                           pd.DataFrame(iris.target,columns=['target'])],axis=1)
iris_stand_df.head()
```

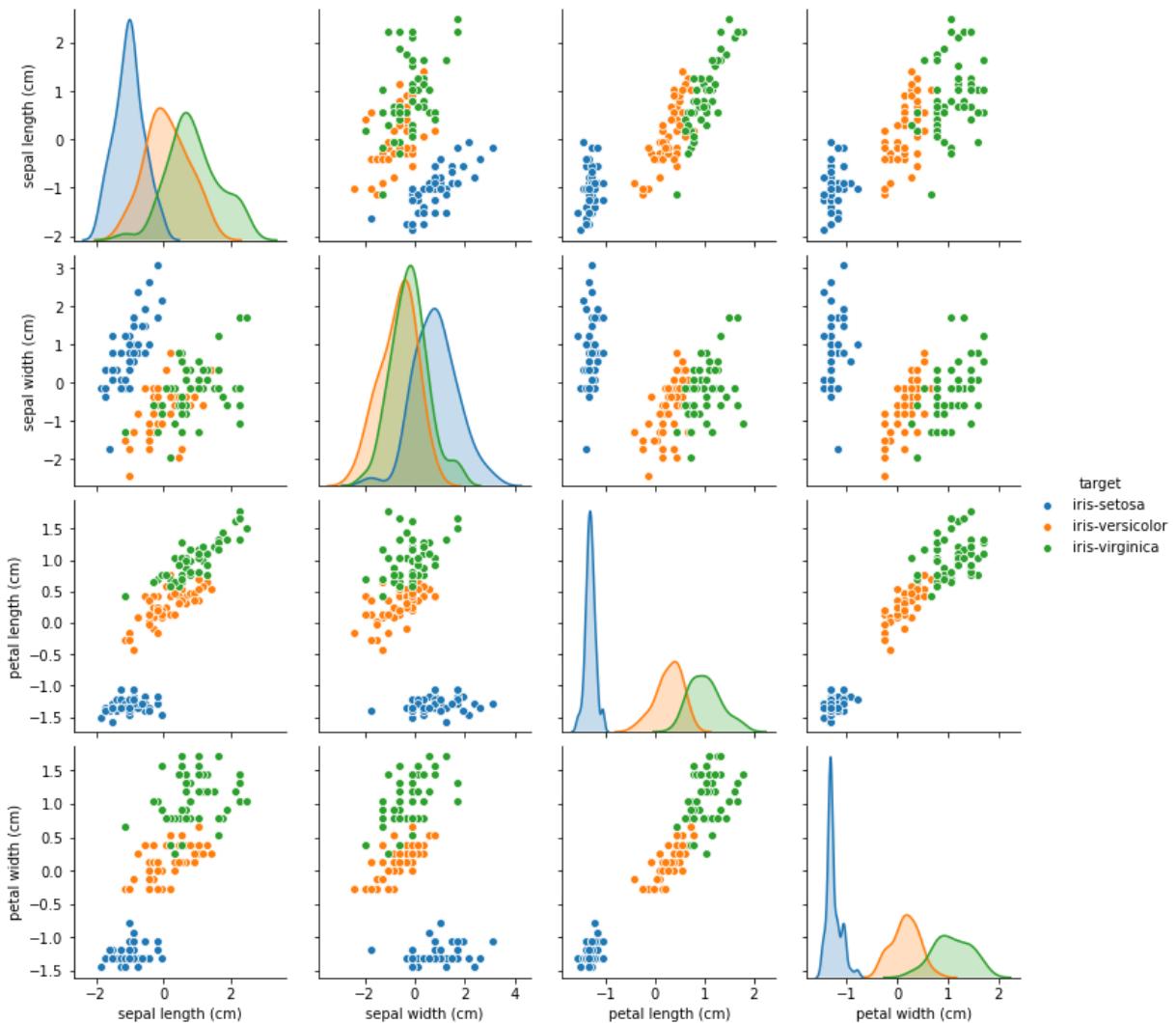
	sepal length (cm)	sepal width (cm)	petal length (cm)	petal width (cm)	target
0	-0.900681	1.019004	-1.340227	-1.315444	0
1	-1.143017	-0.131979	-1.340227	-1.315444	0
2	-1.385353	0.328414	-1.397064	-1.315444	0
3	-1.506521	0.098217	-1.283389	-1.315444	0
4	-1.021849	1.249201	-1.340227	-1.315444	0

```
In [37]: class_label = {0:'iris-setosa',1:'iris-versicolor',2:'iris-virginica'}
```

```
In [38]: iris_stand_df['target'] = iris_stand_df['target'].map(lambda row: class_label[row])
```

```
In [39]: sns.pairplot(data=iris_stand_df,hue='target')
```

```
Out[39]: <seaborn.axisgrid.PairGrid at 0x20f8afe2438>
```



In [28]: `iris_stand_df.shape`

Out[28]: (150, 5)

In [29]: `iris_stand_df.target.value_counts()`

Out[29]:

2	50
1	50
0	50

Name: target, dtype: int64

Multiple_runs

IRIS:CASE-I-Multiple_runs

Running t-SNE on a range of perplexities and iterations with fixed Learning rate and embedding initialization as 'random'

```
In [78]: perplexities = [1, 2, 5, 30, 50, 100, 150]
iterations = [250, 350, 500, 750, 1000, 2500, 3500, 4000, 5000]

for pidx in range(len(perplexities)):
    for idx in range(len(iterations)):
        tsne_iris_1 = TSNE(n_components=2, perplexity=perplexities[pidx], learning_rat
```

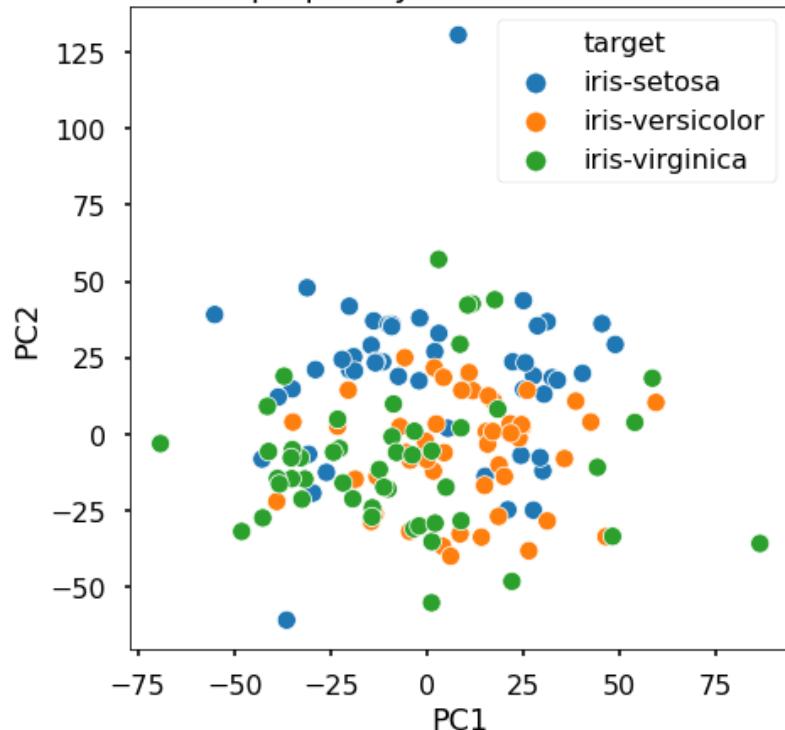
```

TSNE
init='random',n_jobs=-1,random_state=23)
iris1_tsne_pcmps = pd.DataFrame(tsne_iris_1.fit_transform(iris_stand_df.iloc
iris1_tsne_pcmps = pd.concat([iris1_tsne_pcmps,iris_stand_df['target']],axis
with plt.style.context('seaborn-poster'):
    plt.figure(figsize=(7,7))
    sns.scatterplot(data=iris1_tsne_pcmps,x='PC1',y='PC2',hue='target')
    plt.title("t-SNE visualization with perplexity -- {0}, iterations -- {1}

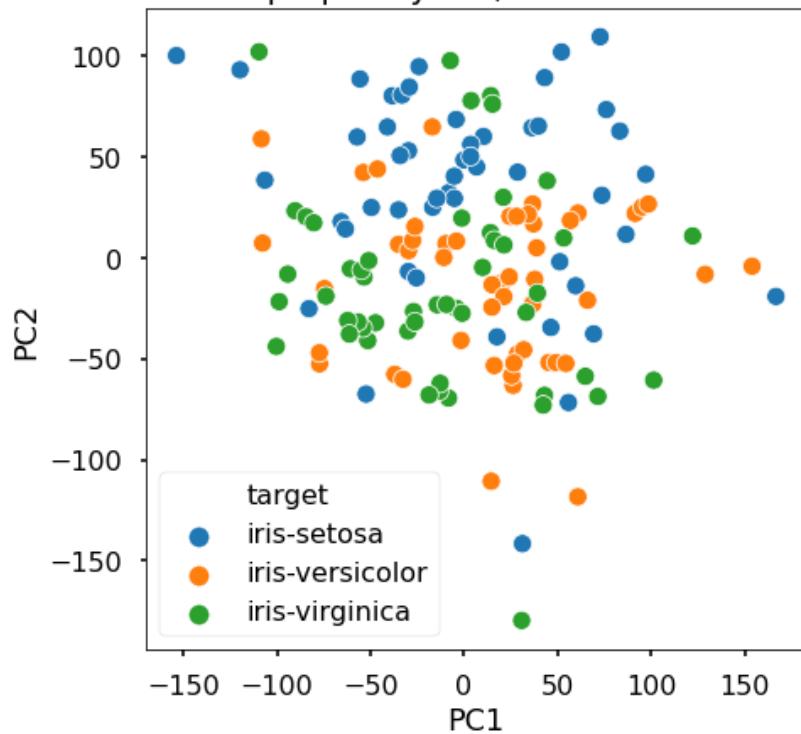
plt.show()

```

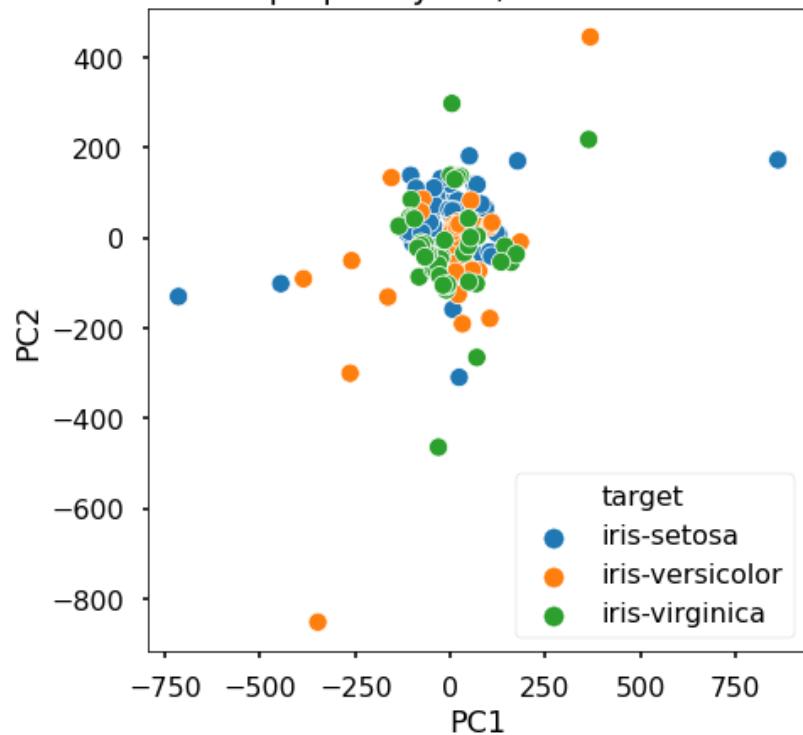
t-SNE visualization with perplexity -- 2, iterations -- 250 and epsilon -- 200



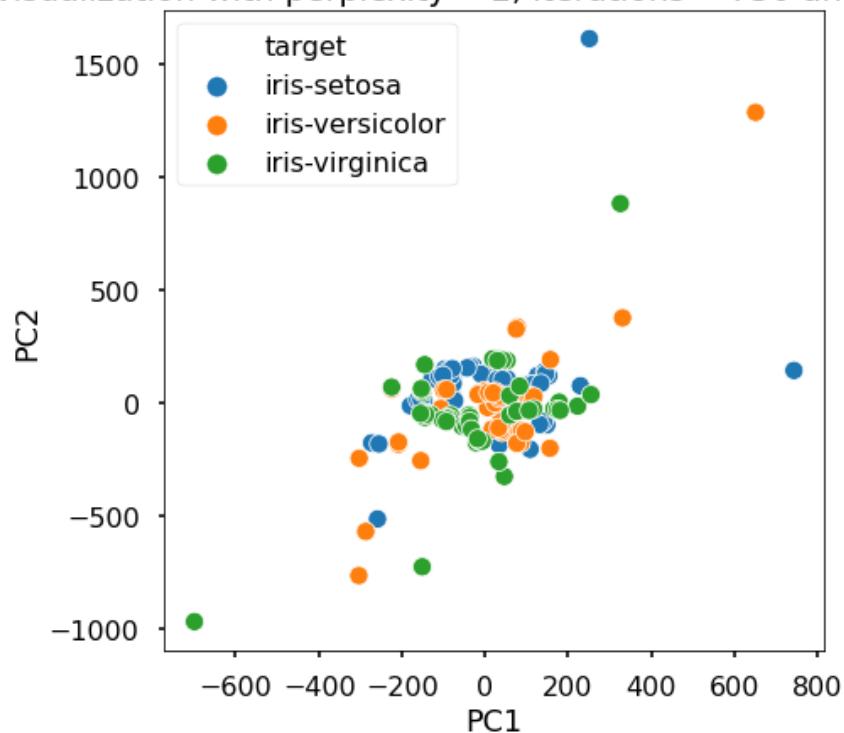
t-SNE visualization with perplexity -- 2, iterations -- 350 and epsilon -- 200



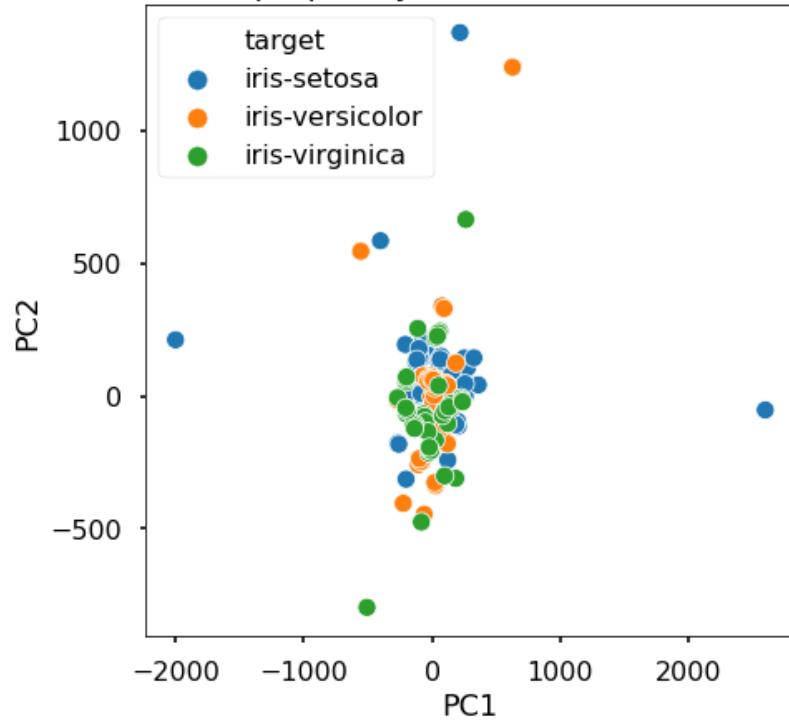
t-SNE visualization with perplexity -- 2, iterations -- 500 and epsilon -- 200



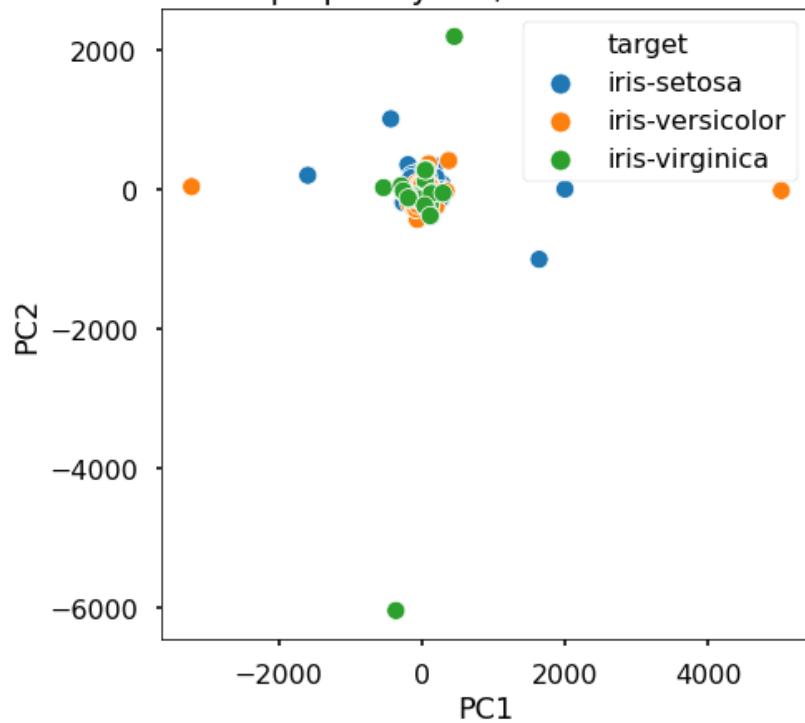
t-SNE visualization with perplexity -- 2, iterations -- 750 and epsilon -- 200



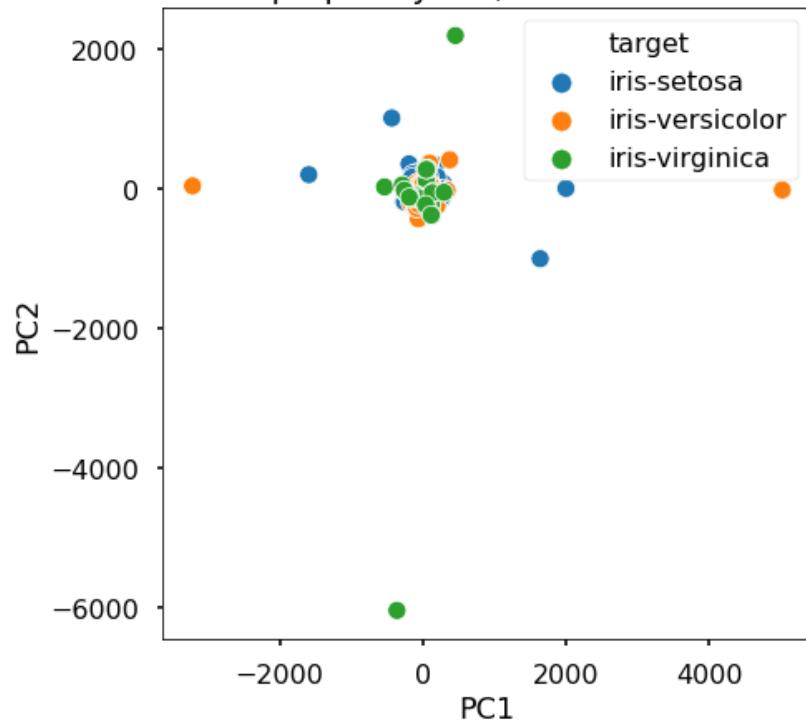
t-SNE visualization with perplexity -- 2, iterations -- 1000 and epsilon -- 200



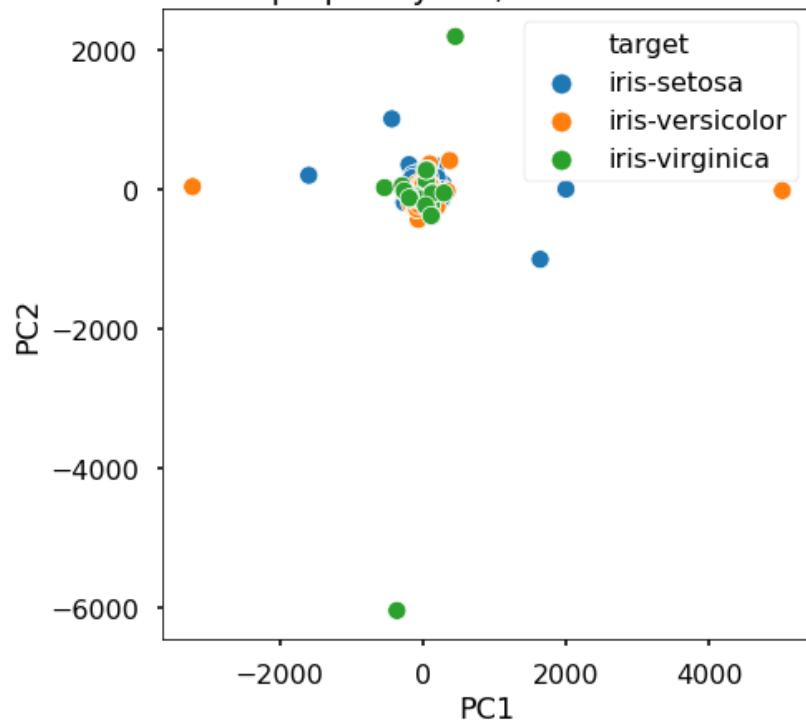
t-SNE visualization with perplexity -- 2, iterations -- 2500 and epsilon -- 200



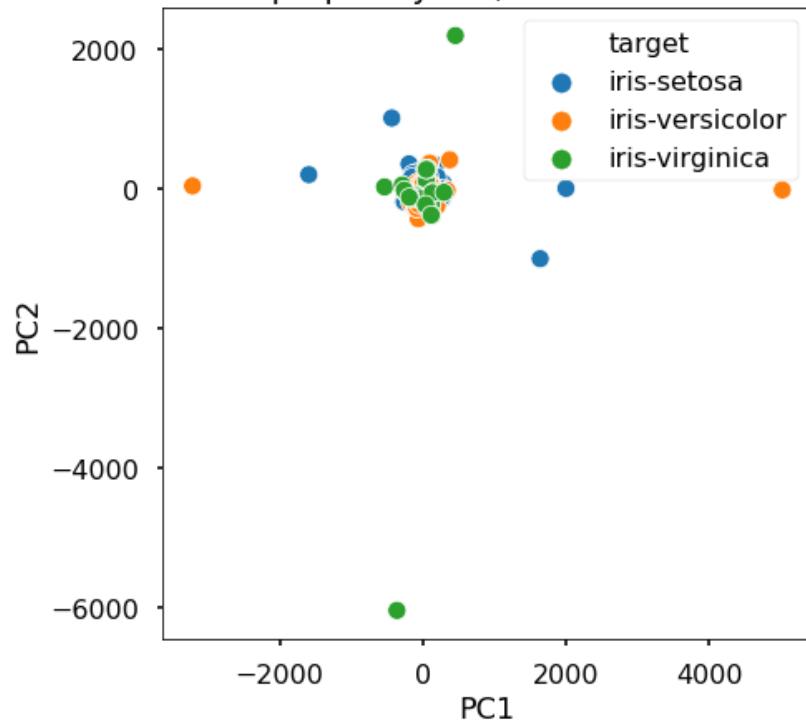
t-SNE visualization with perplexity -- 2, iterations -- 3500 and epsilon -- 200



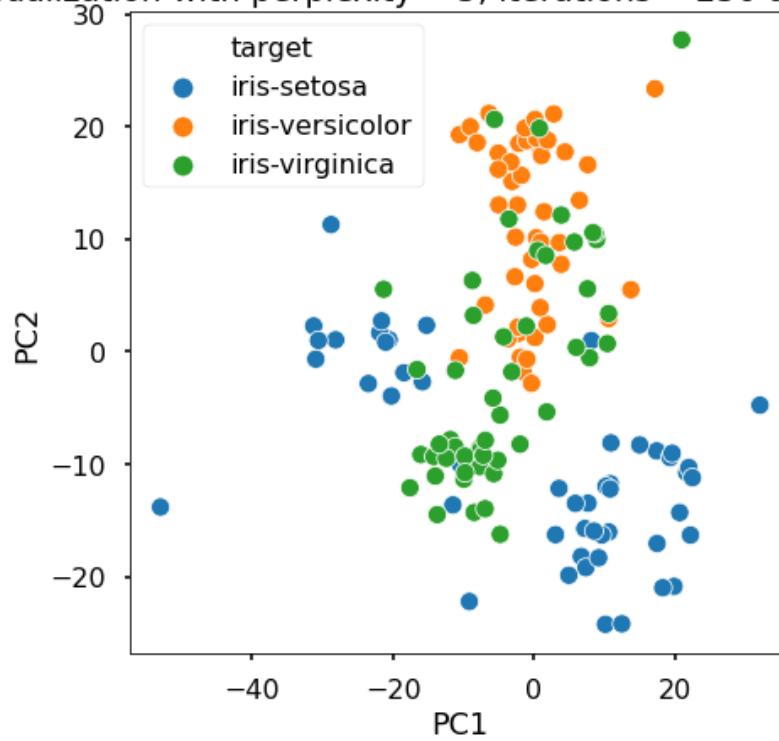
t-SNE visualization with perplexity -- 2, iterations -- 4000 and epsilon -- 200



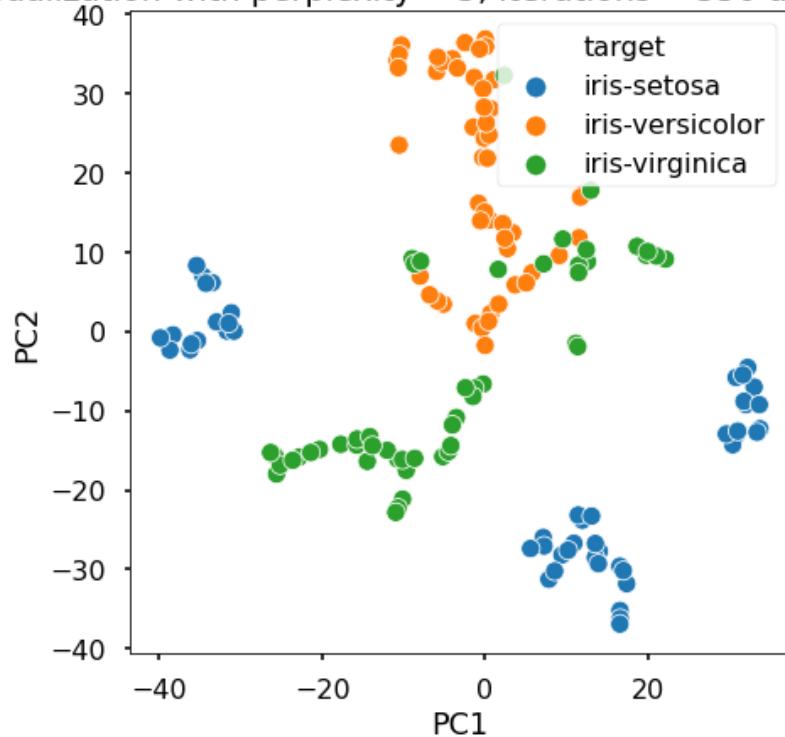
t-SNE visualization with perplexity -- 2, iterations -- 5000 and epsilon -- 200



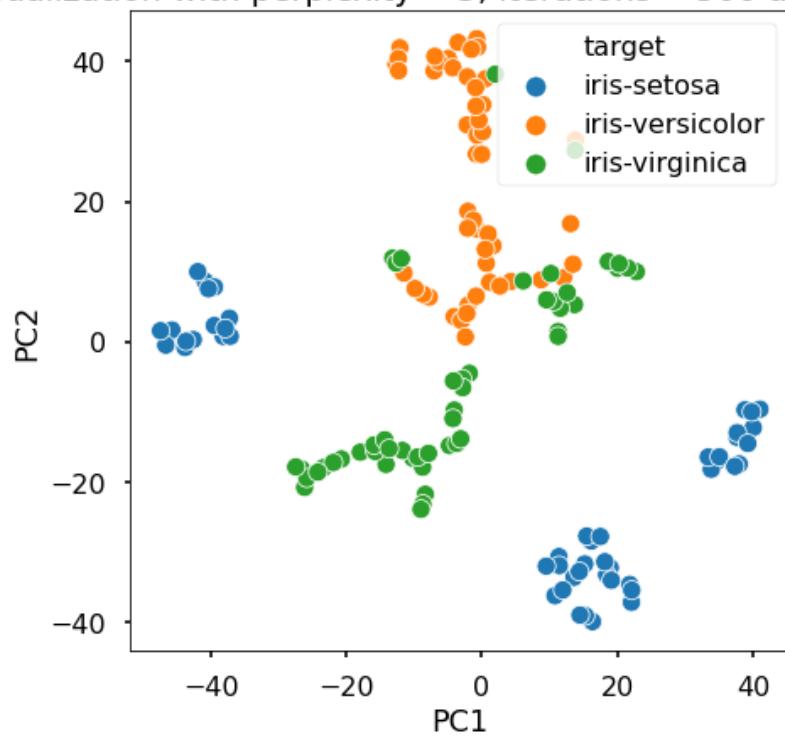
t-SNE visualization with perplexity -- 5, iterations -- 250 and epsilon -- 200



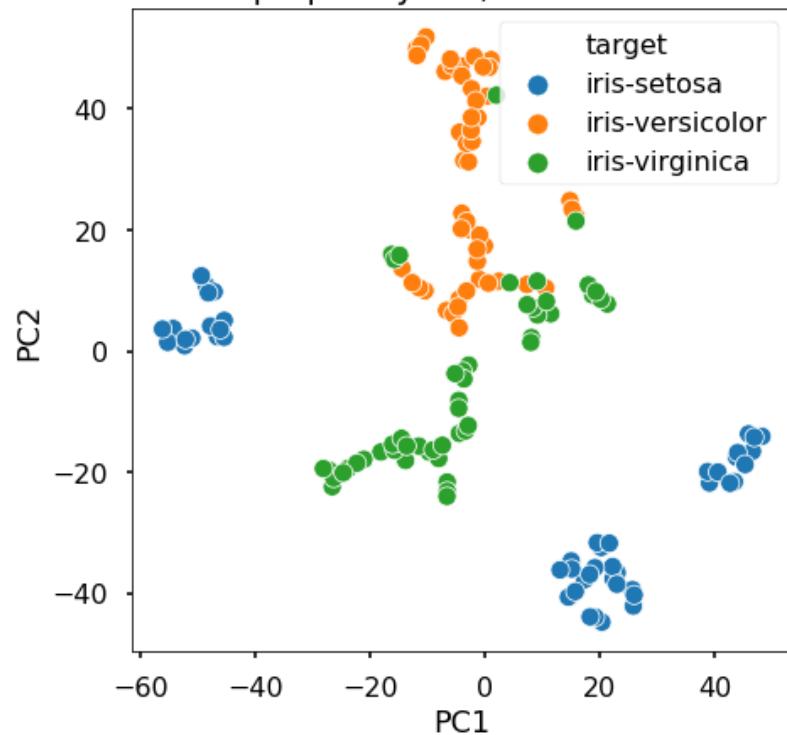
t-SNE visualization with perplexity -- 5, iterations -- 350 and epsilon -- 200



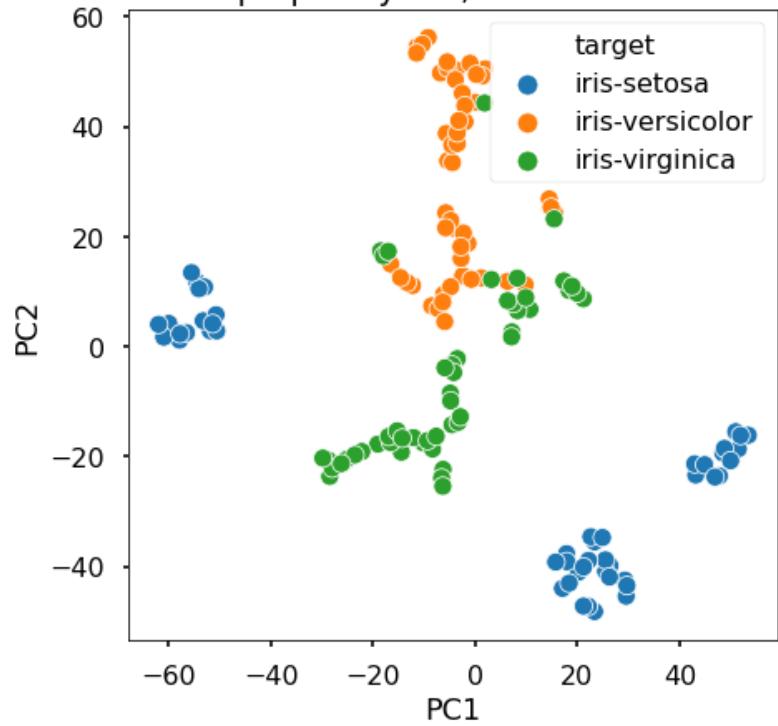
t-SNE visualization with perplexity -- 5, iterations -- 500 and epsilon -- 200



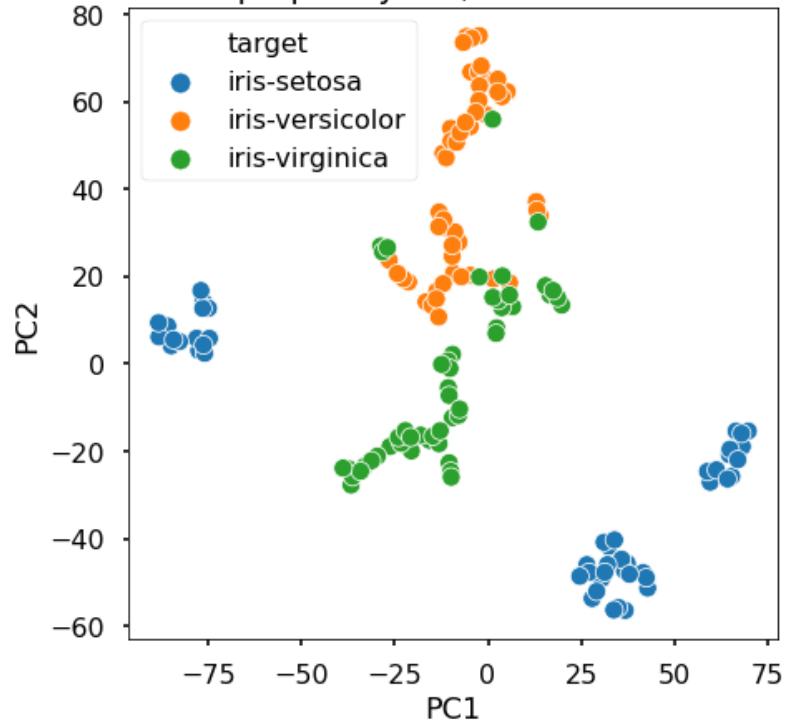
t-SNE visualization with perplexity -- 5, iterations -- 750 and epsilon -- 200



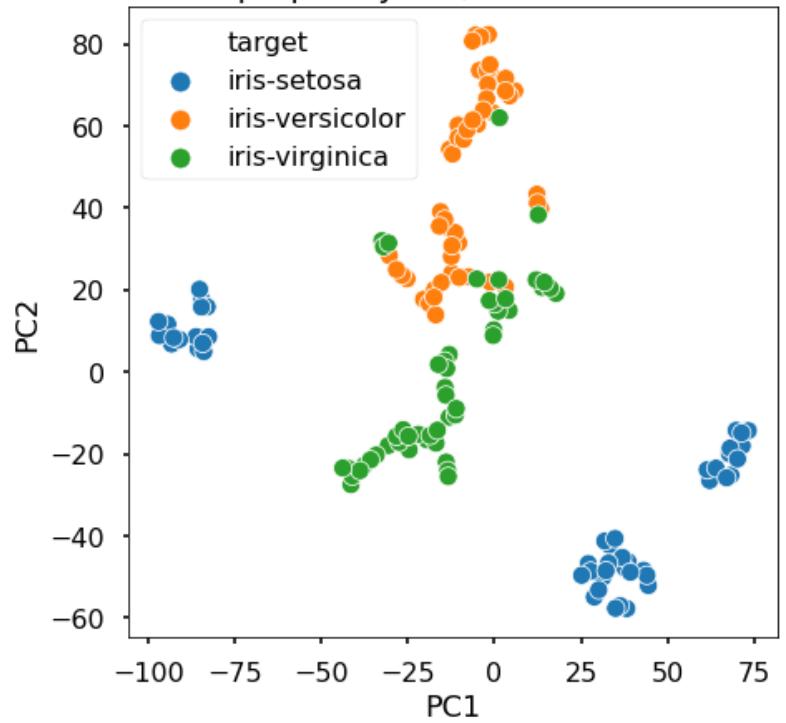
t-SNE visualization with perplexity -- 5, iterations -- 1000 and epsilon -- 200



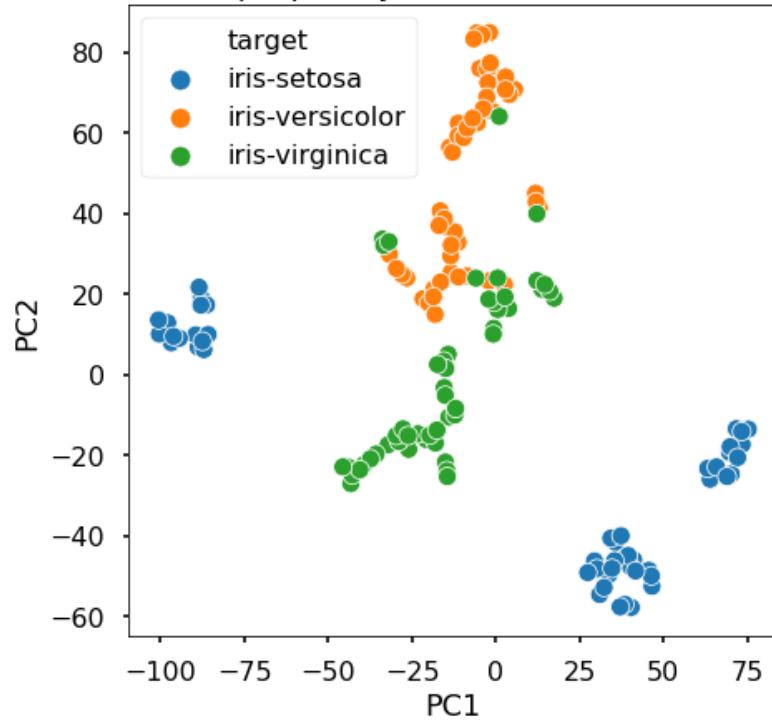
t-SNE visualization with perplexity -- 5, iterations -- 2500 and epsilon -- 200



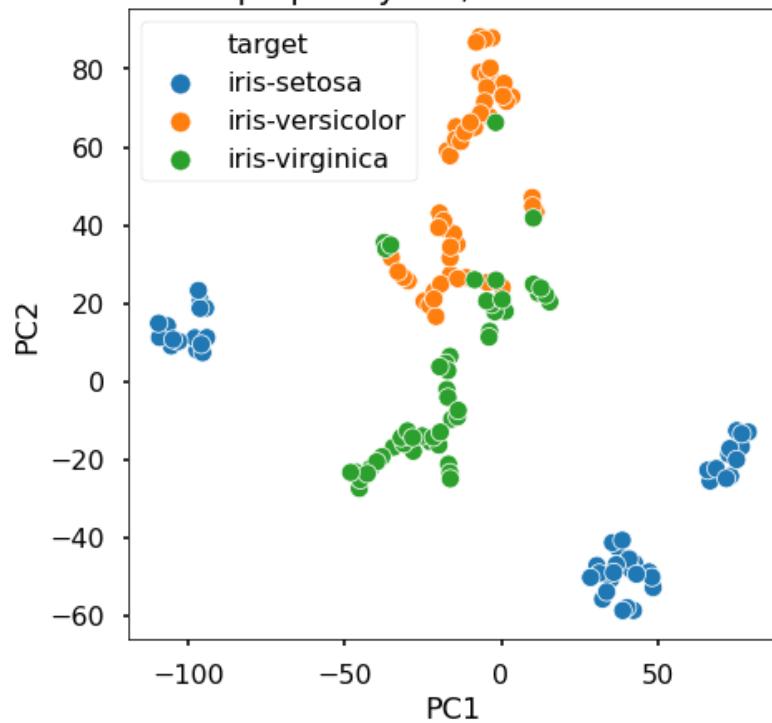
t-SNE visualization with perplexity -- 5, iterations -- 3500 and epsilon -- 200



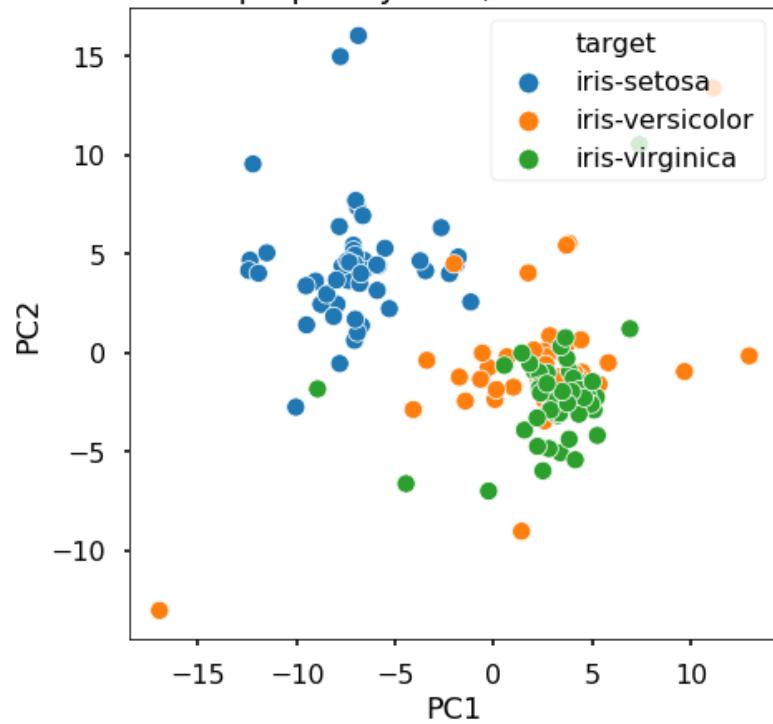
t-SNE visualization with perplexity -- 5, iterations -- 4000 and epsilon -- 200



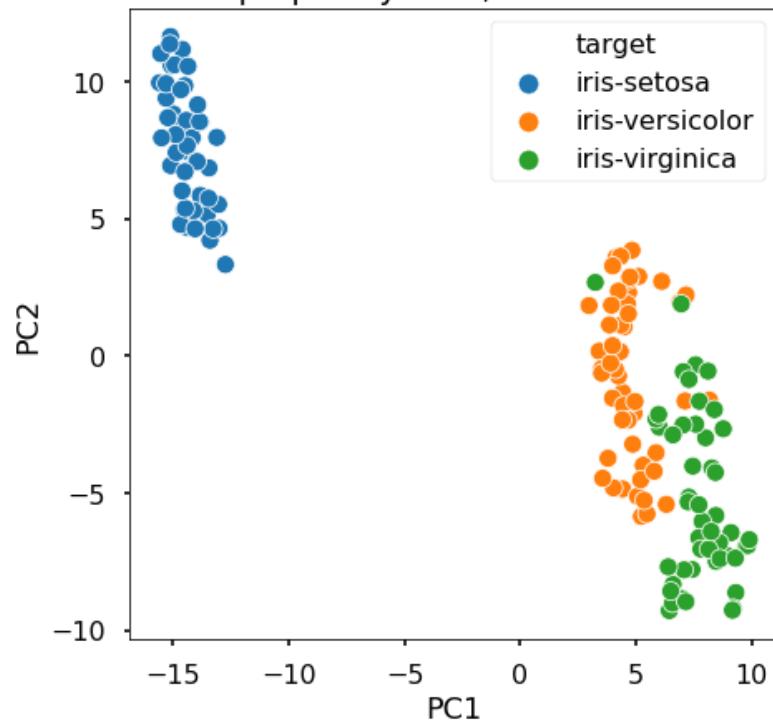
t-SNE visualization with perplexity -- 5, iterations -- 5000 and epsilon -- 200



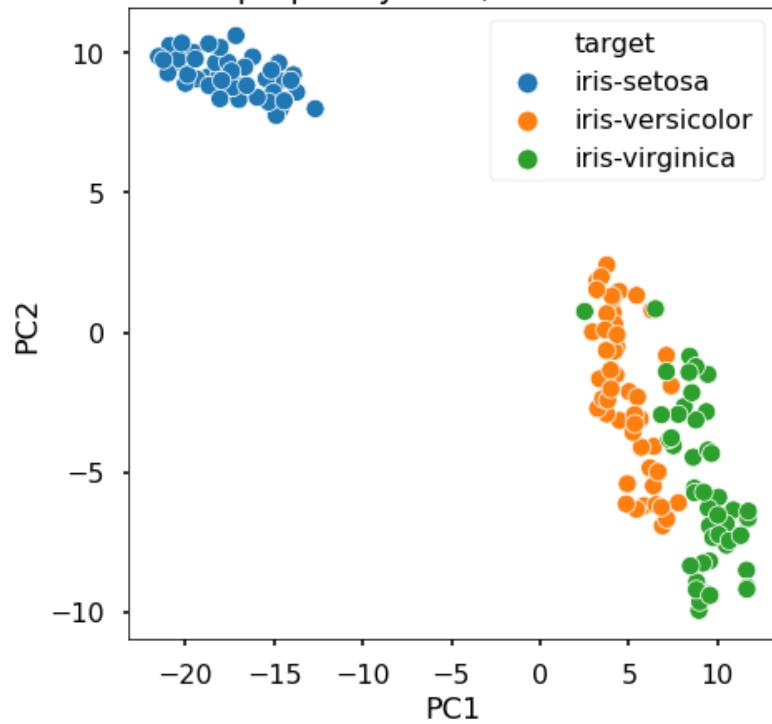
t-SNE visualization with perplexity -- 30, iterations -- 250 and epsilon -- 200



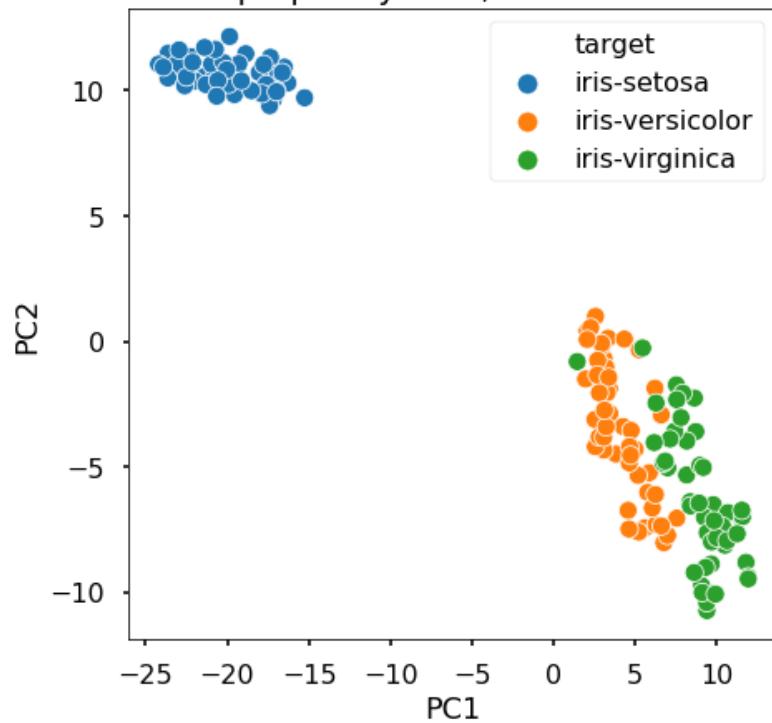
t-SNE visualization with perplexity -- 30, iterations -- 350 and epsilon -- 200



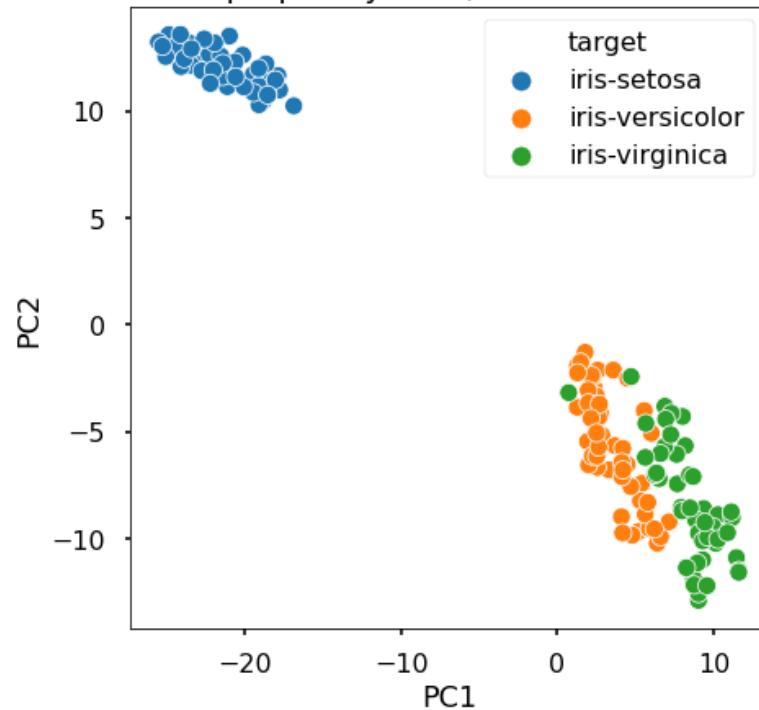
t-SNE visualization with perplexity -- 30, iterations -- 500 and epsilon -- 200



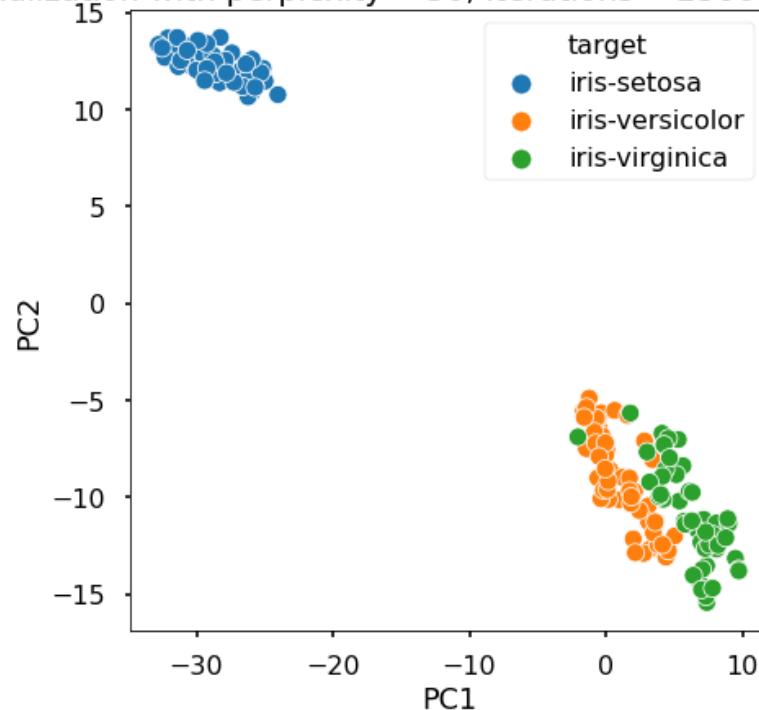
t-SNE visualization with perplexity -- 30, iterations -- 750 and epsilon -- 200



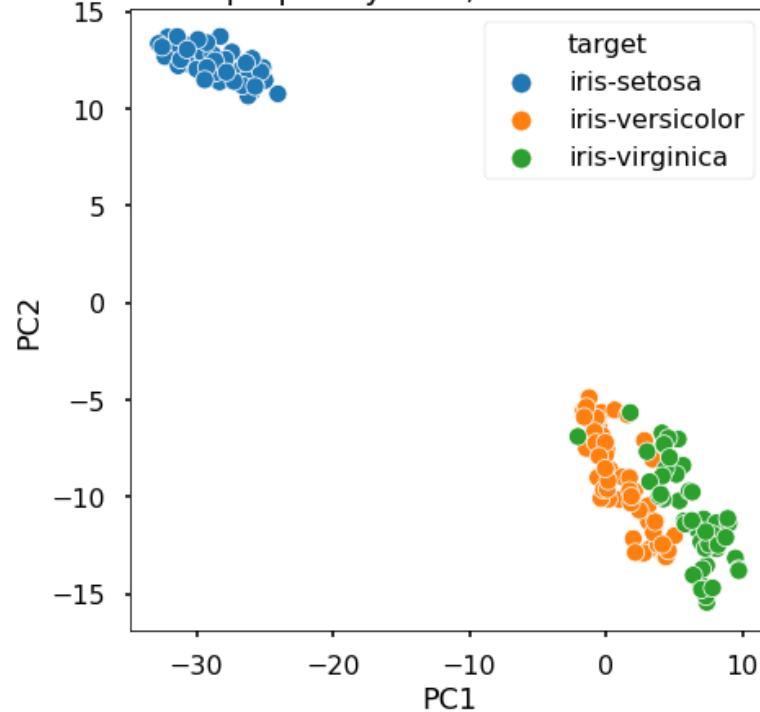
t-SNE visualization with perplexity -- 30, iterations -- 1000 and epsilon -- 200



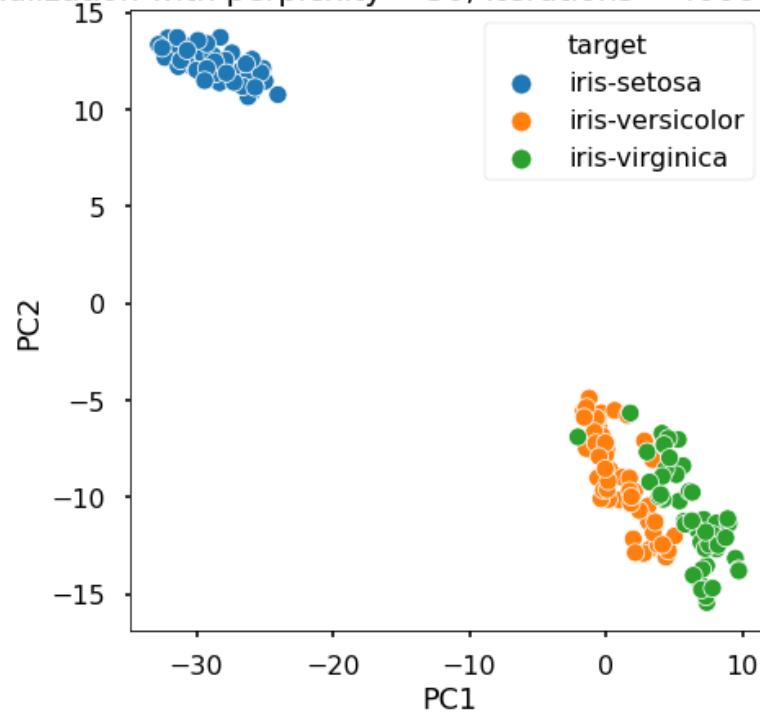
t-SNE visualization with perplexity -- 30, iterations -- 2500 and epsilon -- 200



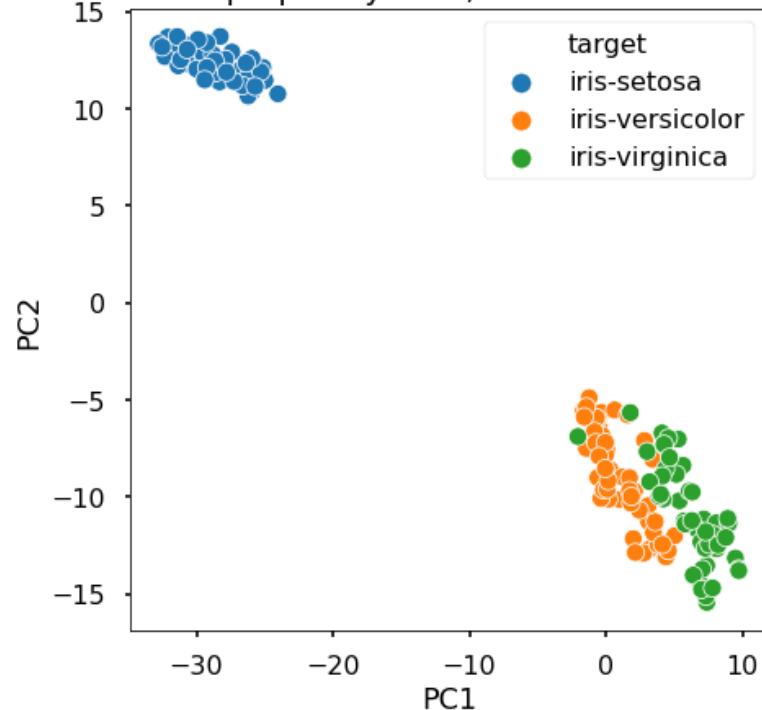
t-SNE visualization with perplexity -- 30, iterations -- 3500 and epsilon -- 200



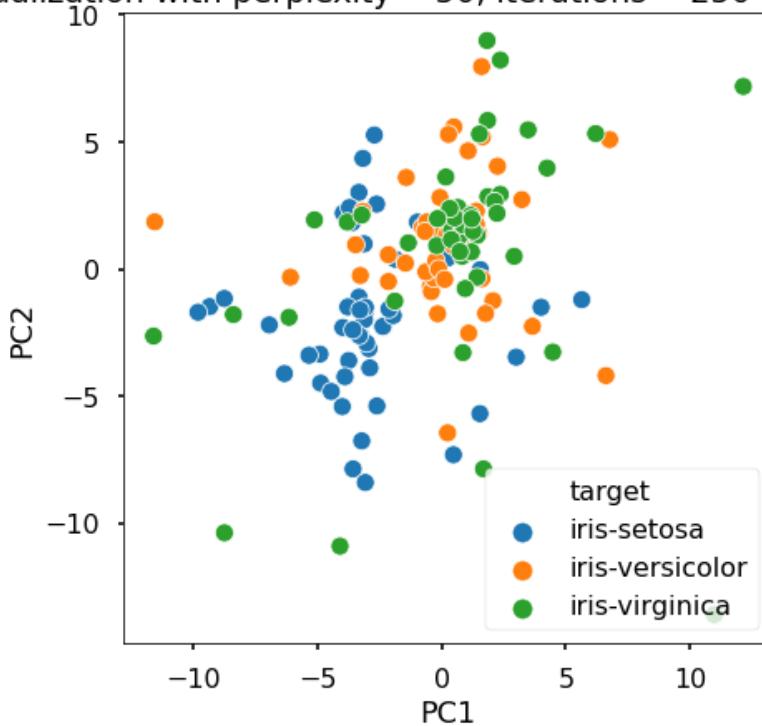
t-SNE visualization with perplexity -- 30, iterations -- 4000 and epsilon -- 200



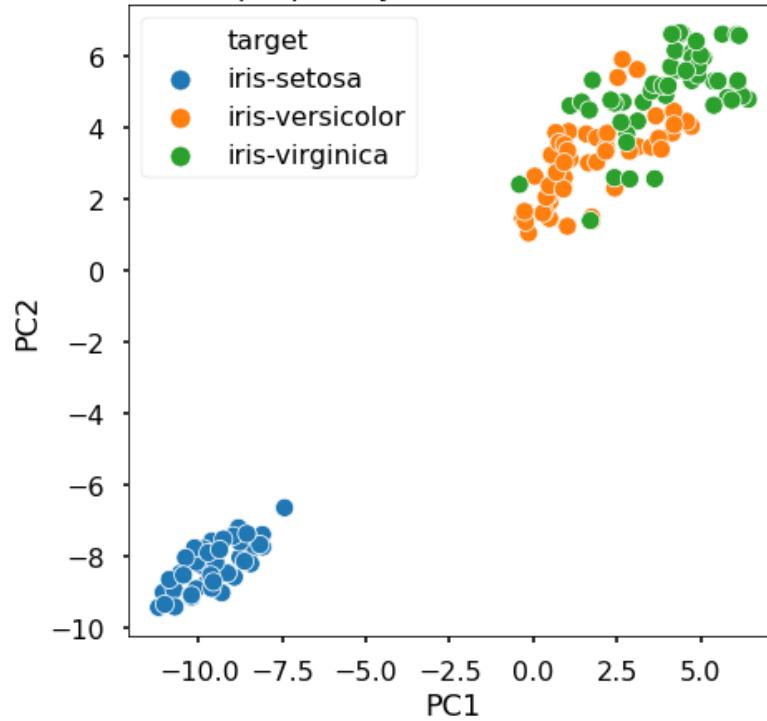
t-SNE visualization with perplexity -- 30, iterations -- 5000 and epsilon -- 200



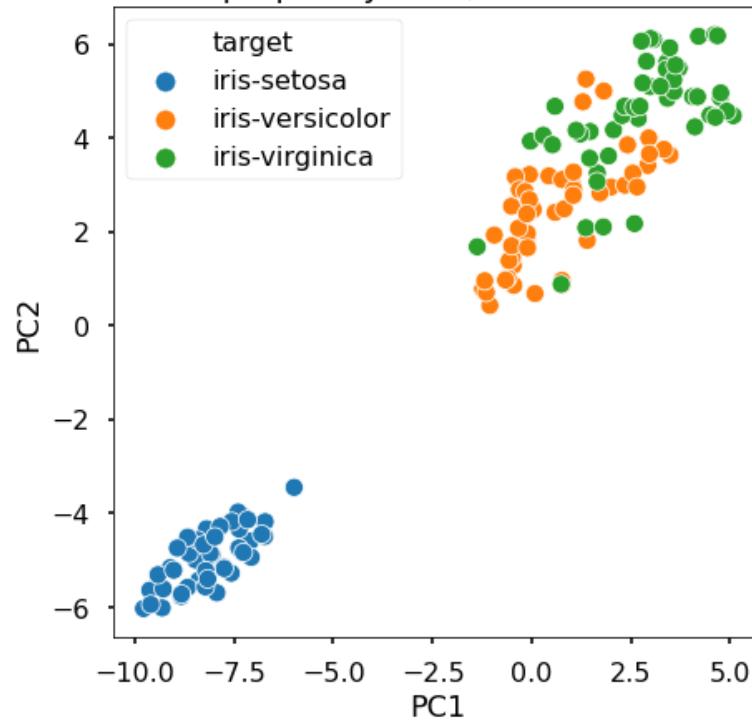
t-SNE visualization with perplexity -- 50, iterations -- 250 and epsilon -- 200



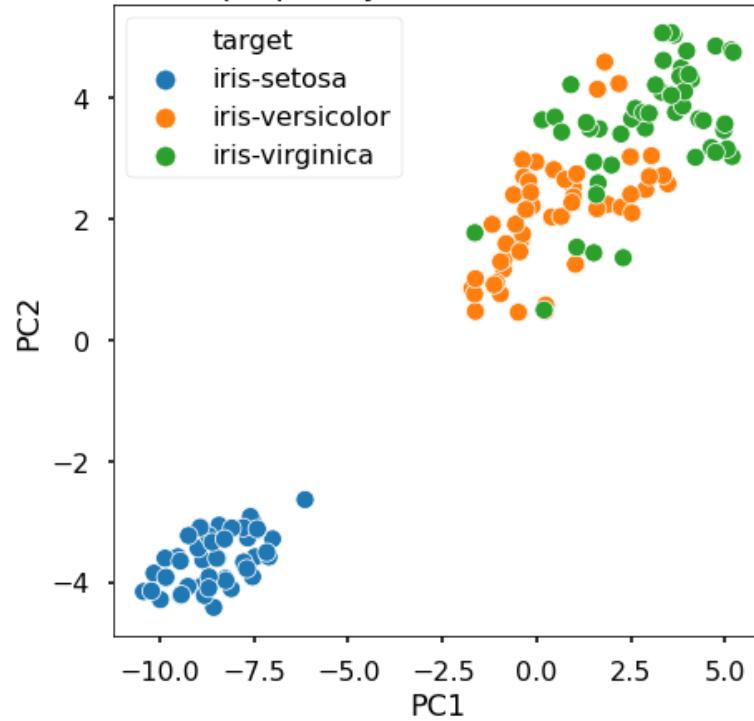
t-SNE visualization with perplexity -- 50, iterations -- 350 and epsilon -- 200



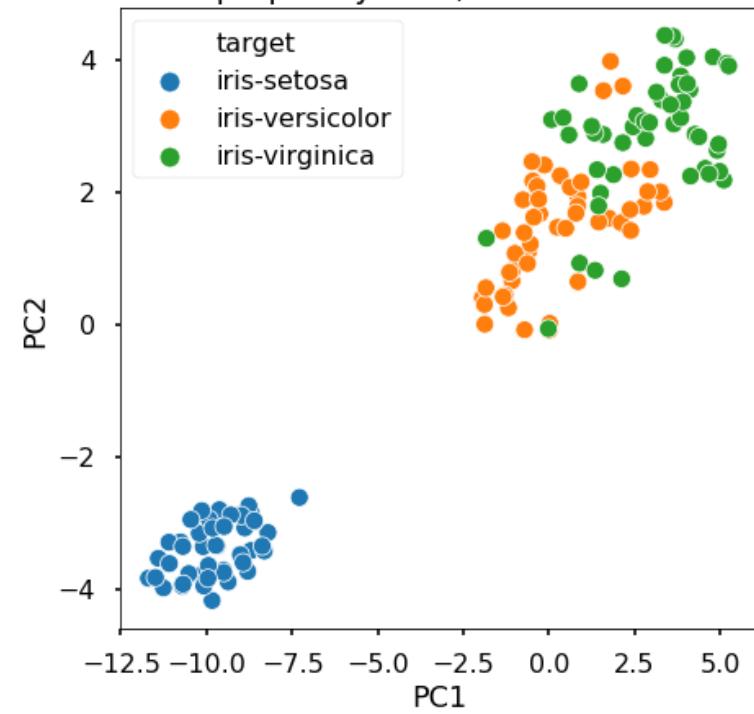
t-SNE visualization with perplexity -- 50, iterations -- 500 and epsilon -- 200



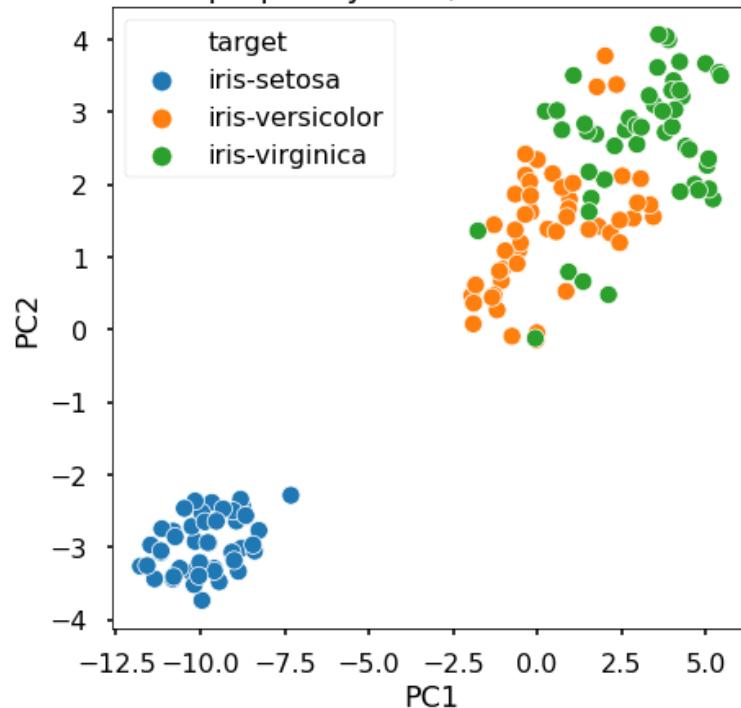
t-SNE visualization with perplexity -- 50, iterations -- 750 and epsilon -- 200



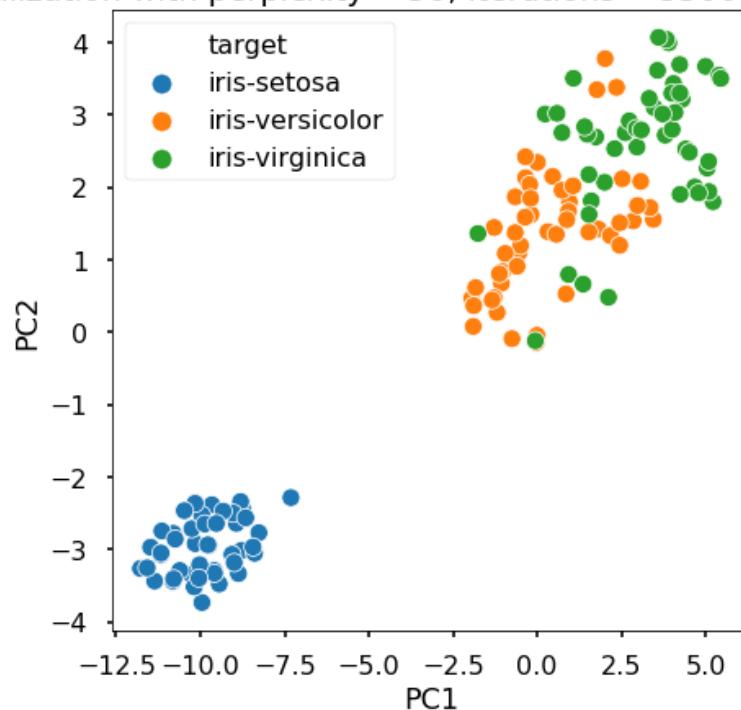
t-SNE visualization with perplexity -- 50, iterations -- 1000 and epsilon -- 200



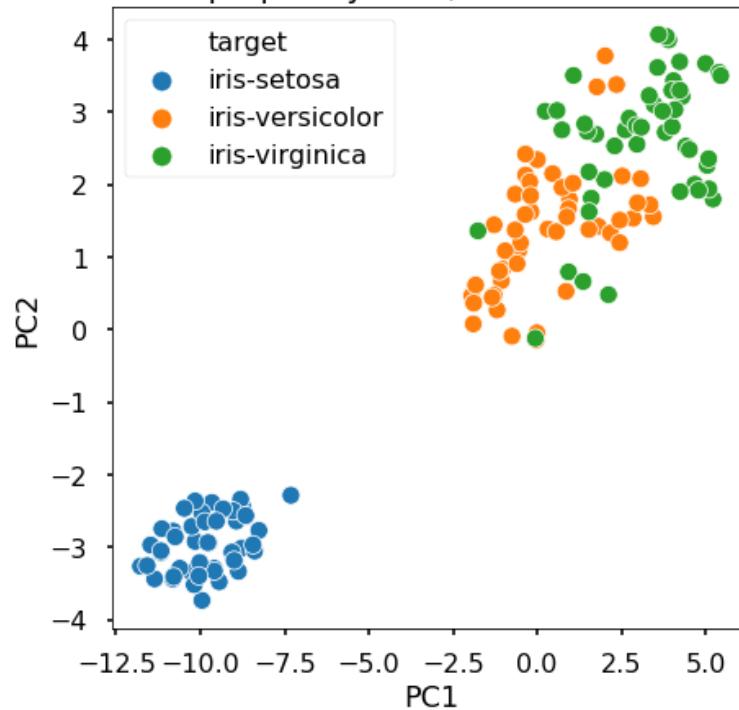
t-SNE visualization with perplexity -- 50, iterations -- 2500 and epsilon -- 200



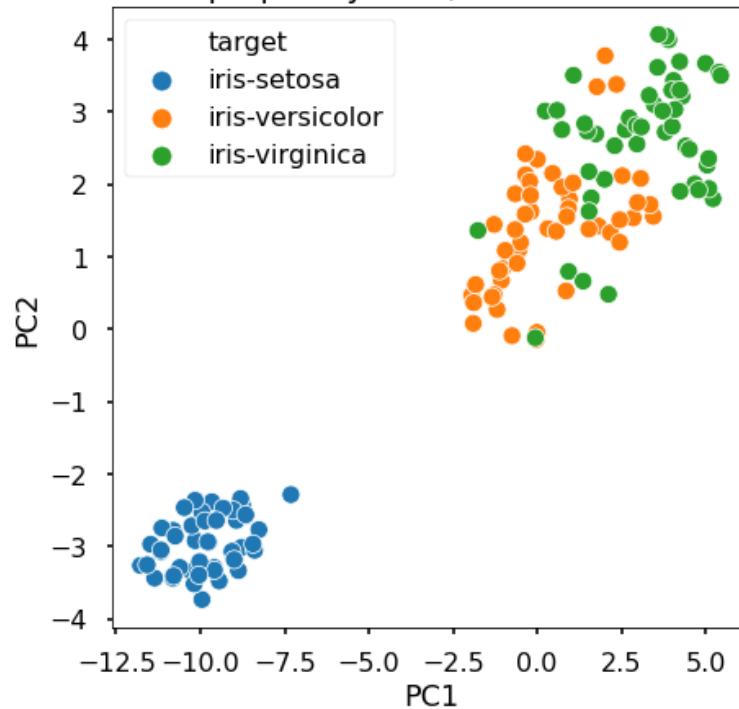
t-SNE visualization with perplexity -- 50, iterations -- 3500 and epsilon -- 200



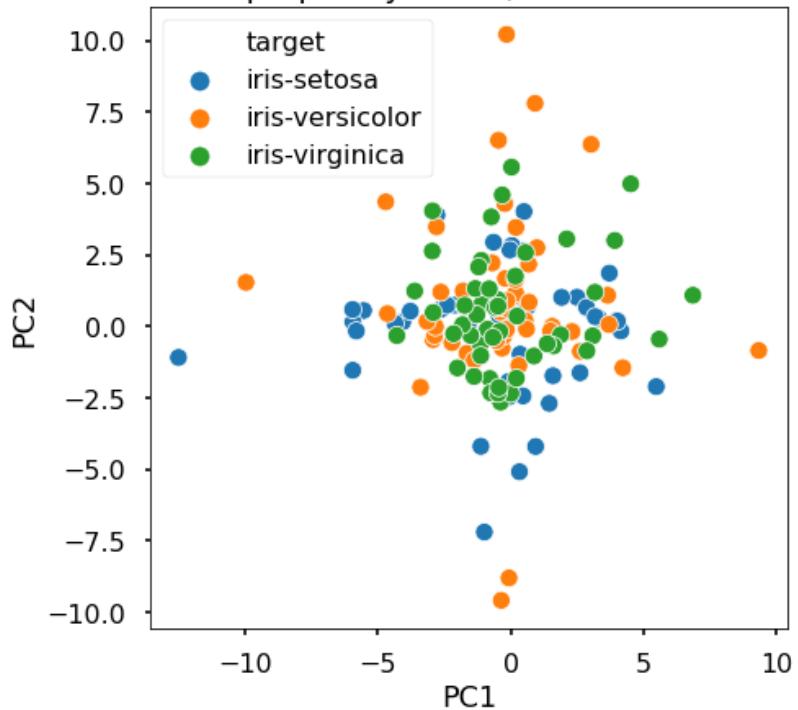
t-SNE visualization with perplexity -- 50, iterations -- 4000 and epsilon -- 200



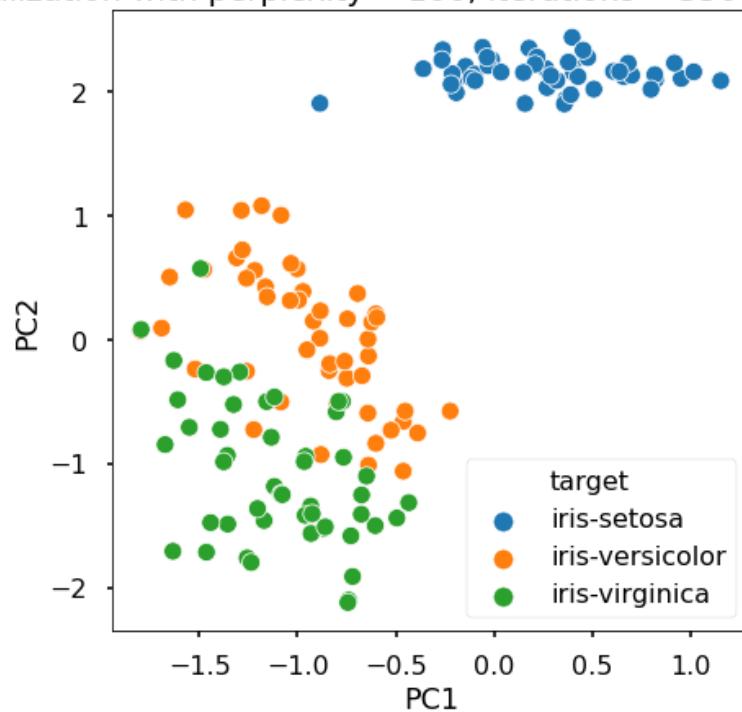
t-SNE visualization with perplexity -- 50, iterations -- 5000 and epsilon -- 200



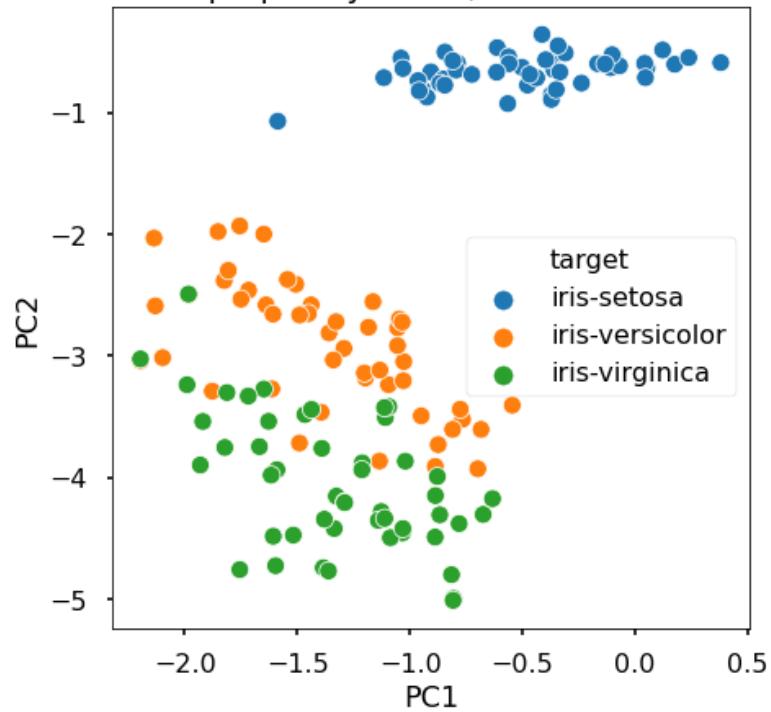
t-SNE visualization with perplexity -- 100, iterations -- 250 and epsilon -- 200



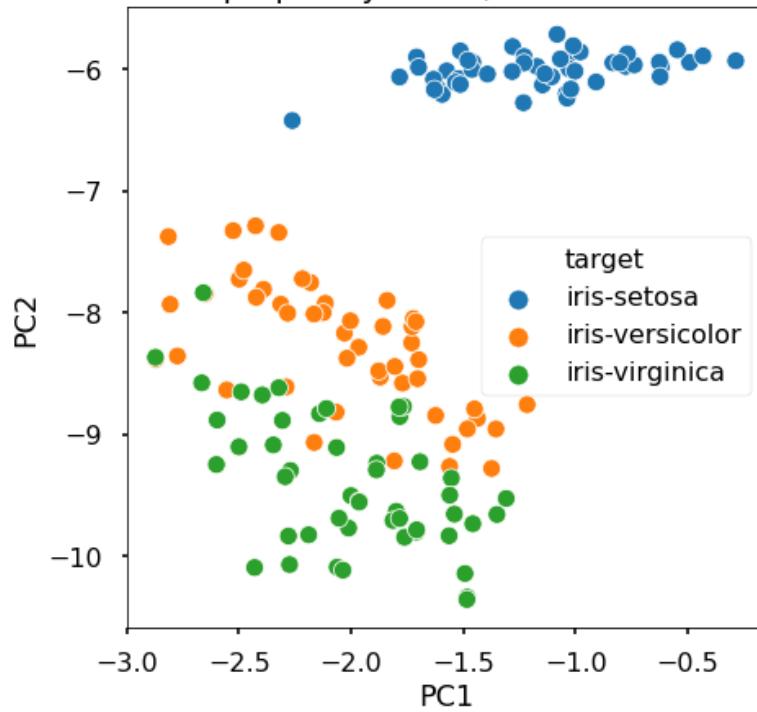
t-SNE visualization with perplexity -- 100, iterations -- 350 and epsilon -- 200



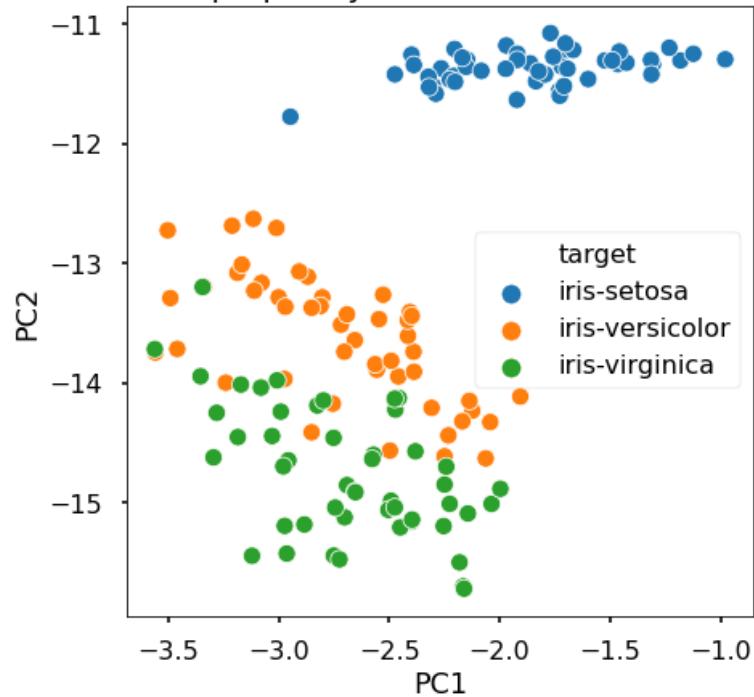
t-SNE visualization with perplexity -- 100, iterations -- 500 and epsilon -- 200



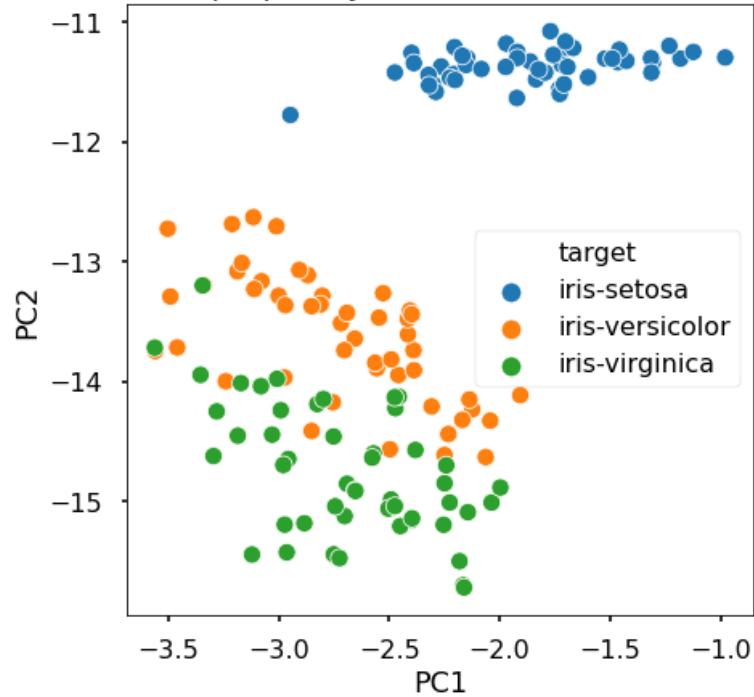
t-SNE visualization with perplexity -- 100, iterations -- 750 and epsilon -- 200



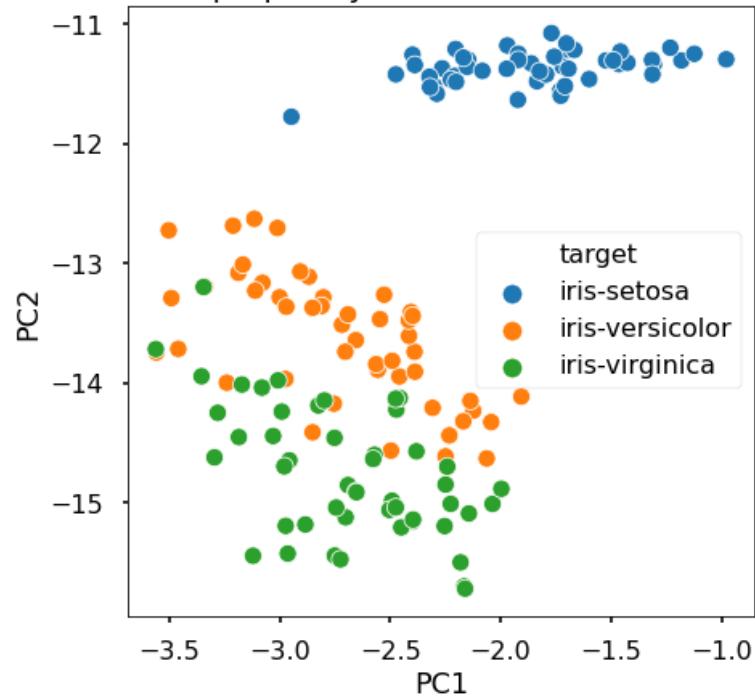
t-SNE visualization with perplexity -- 100, iterations -- 1000 and epsilon -- 200



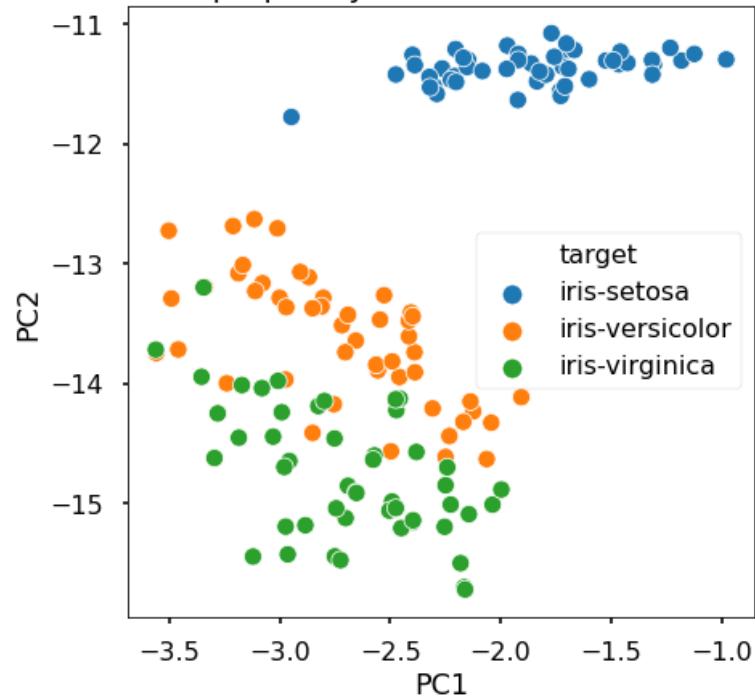
t-SNE visualization with perplexity -- 100, iterations -- 2500 and epsilon -- 200



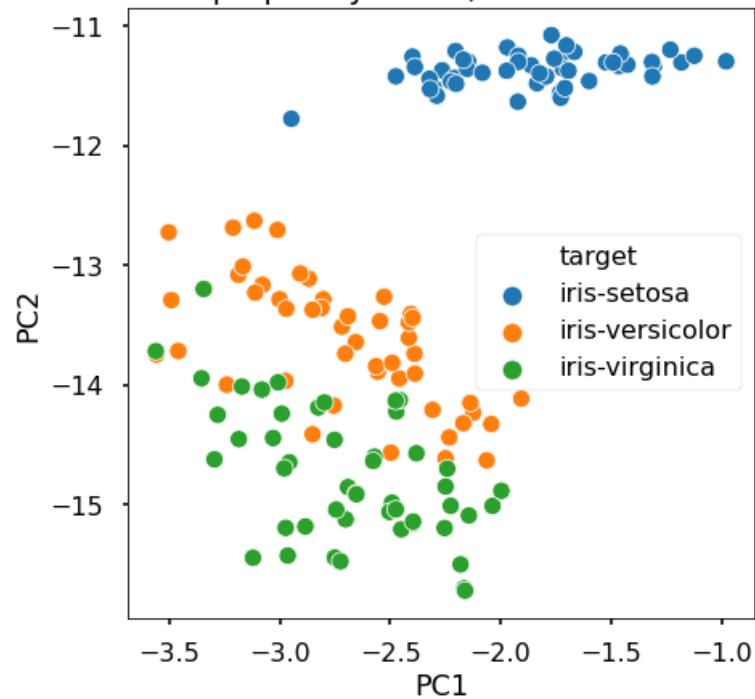
t-SNE visualization with perplexity -- 100, iterations -- 3500 and epsilon -- 200



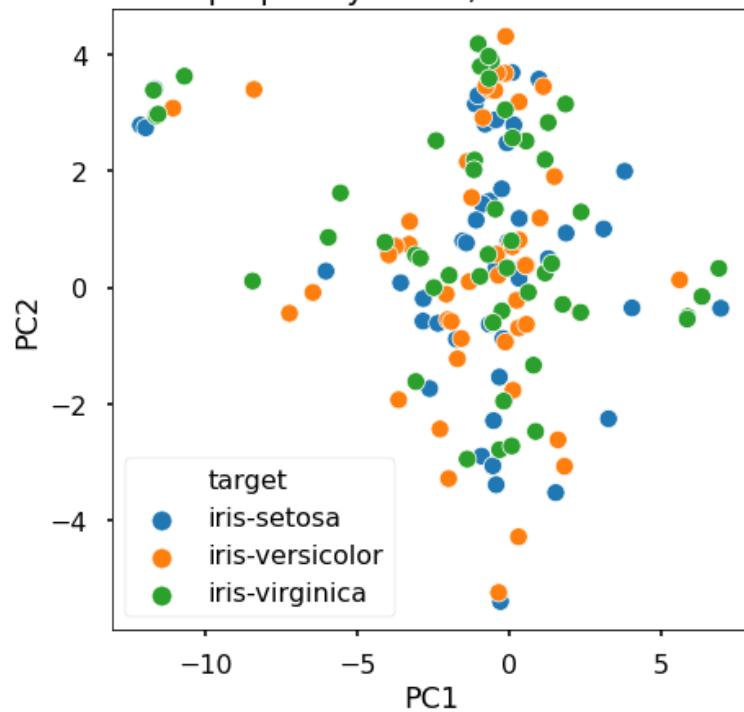
t-SNE visualization with perplexity -- 100, iterations -- 4000 and epsilon -- 200



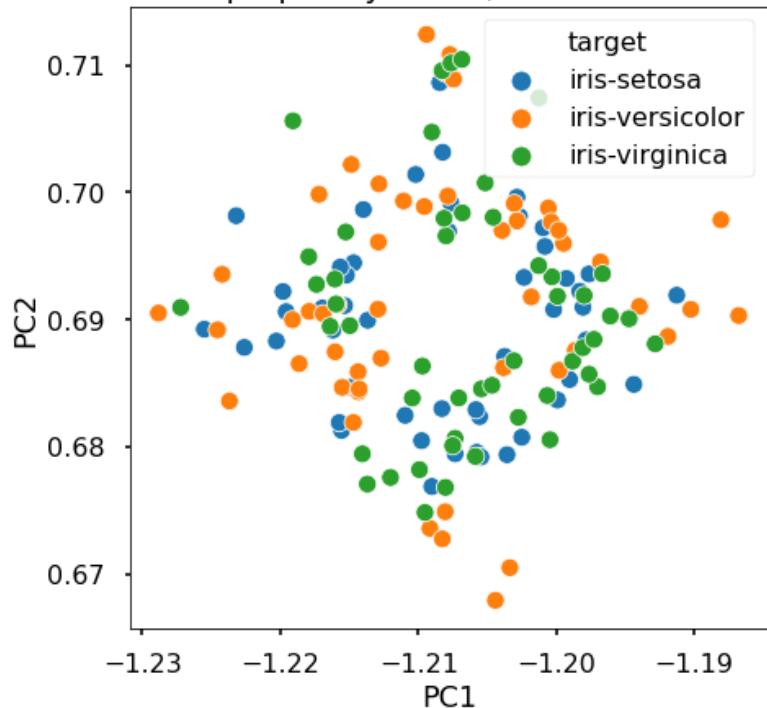
t-SNE visualization with perplexity -- 100, iterations -- 5000 and epsilon -- 200



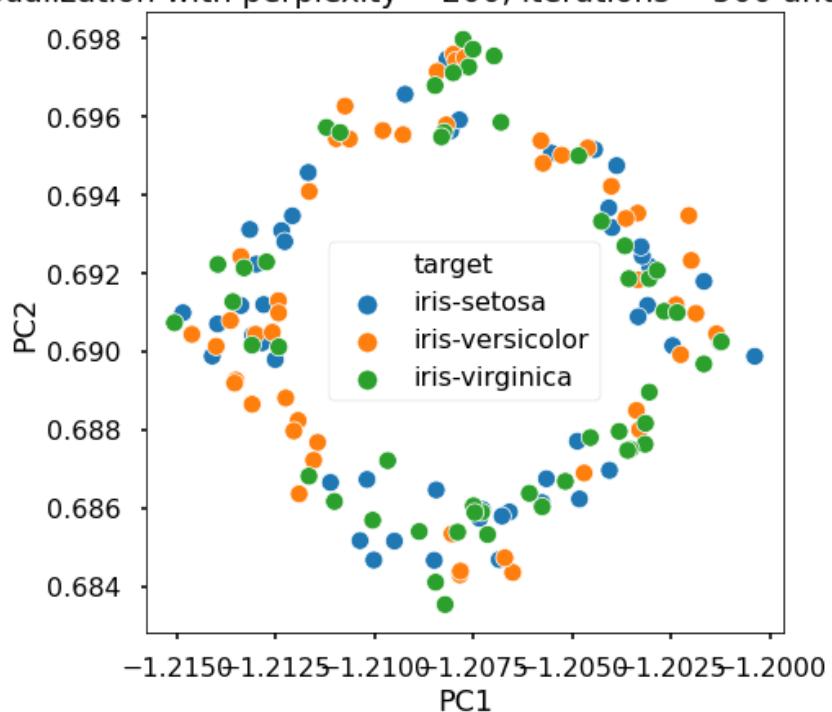
t-SNE visualization with perplexity -- 200, iterations -- 250 and epsilon -- 200



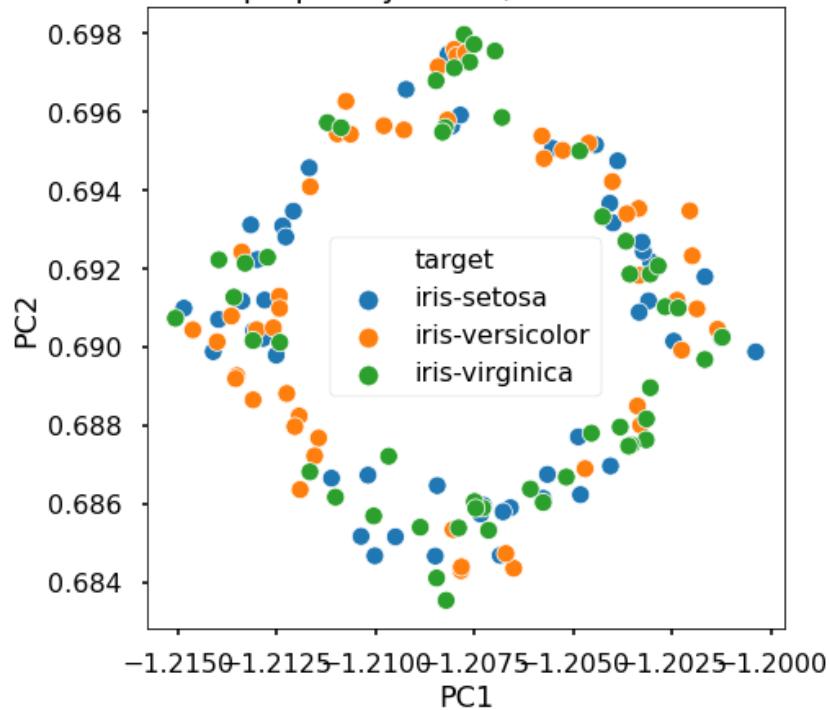
t-SNE visualization with perplexity -- 200, iterations -- 350 and epsilon -- 200



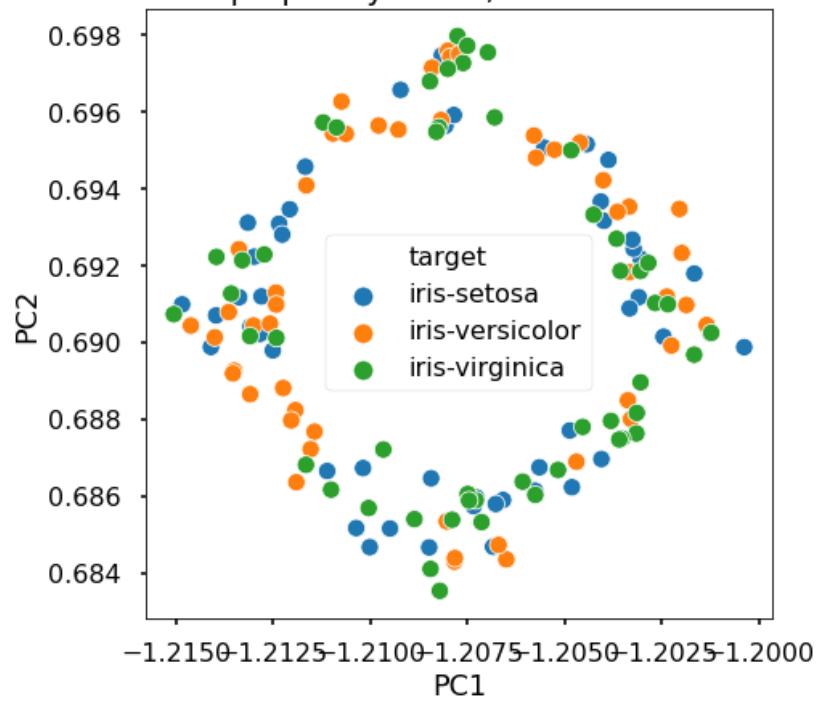
t-SNE visualization with perplexity -- 200, iterations -- 500 and epsilon -- 200



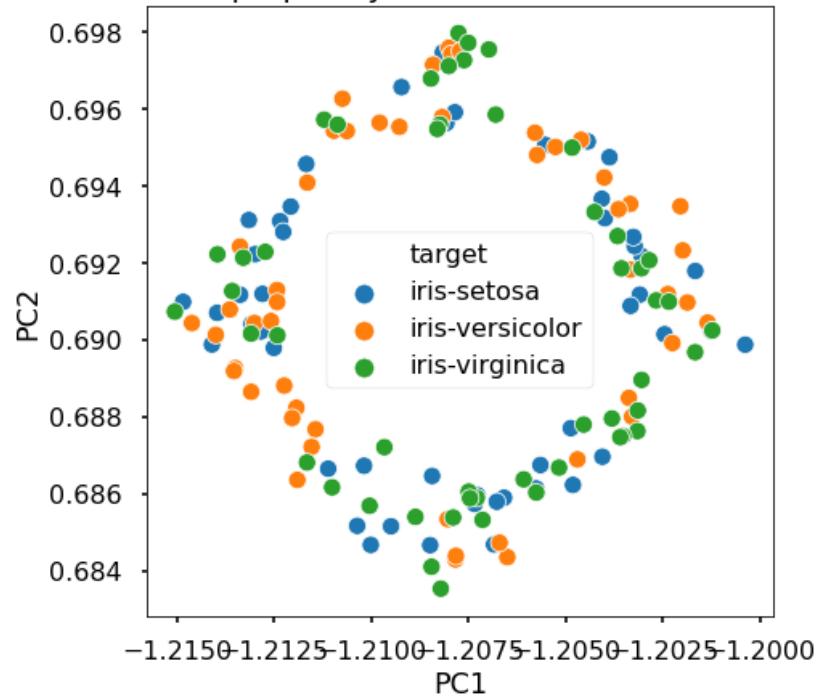
t-SNE visualization with perplexity -- 200, iterations -- 750 and epsilon -- 200



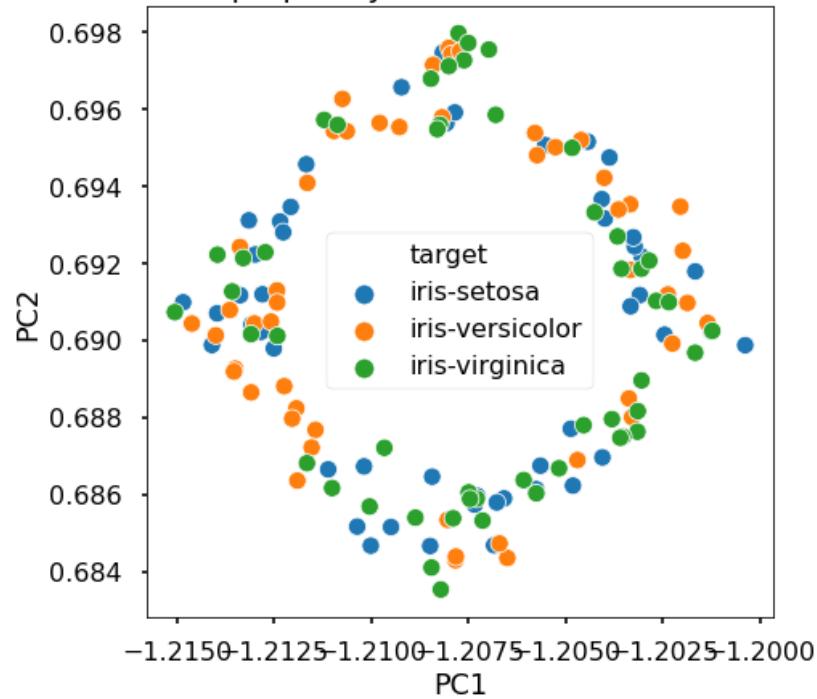
t-SNE visualization with perplexity -- 200, iterations -- 1000 and epsilon -- 200



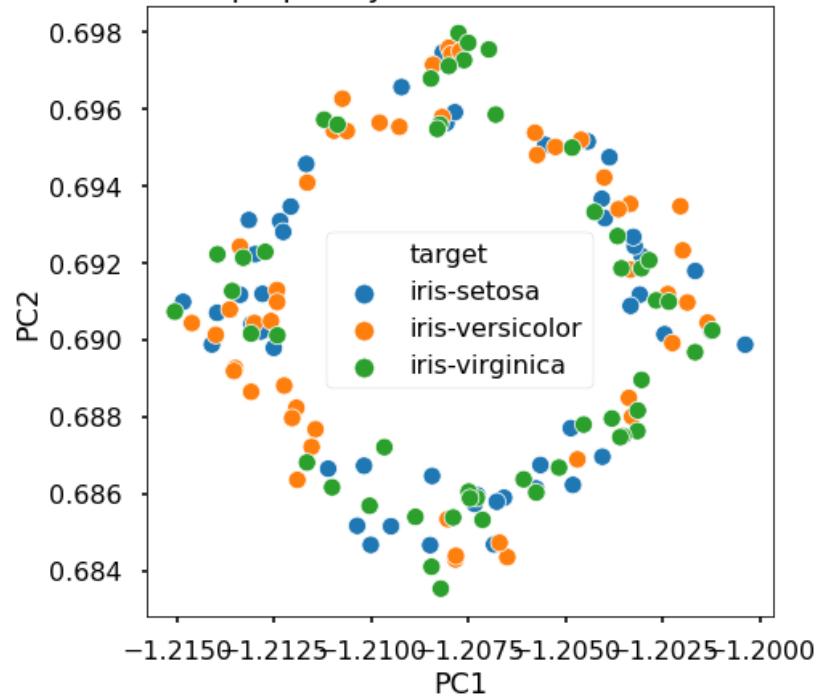
t-SNE visualization with perplexity -- 200, iterations -- 2500 and epsilon -- 200



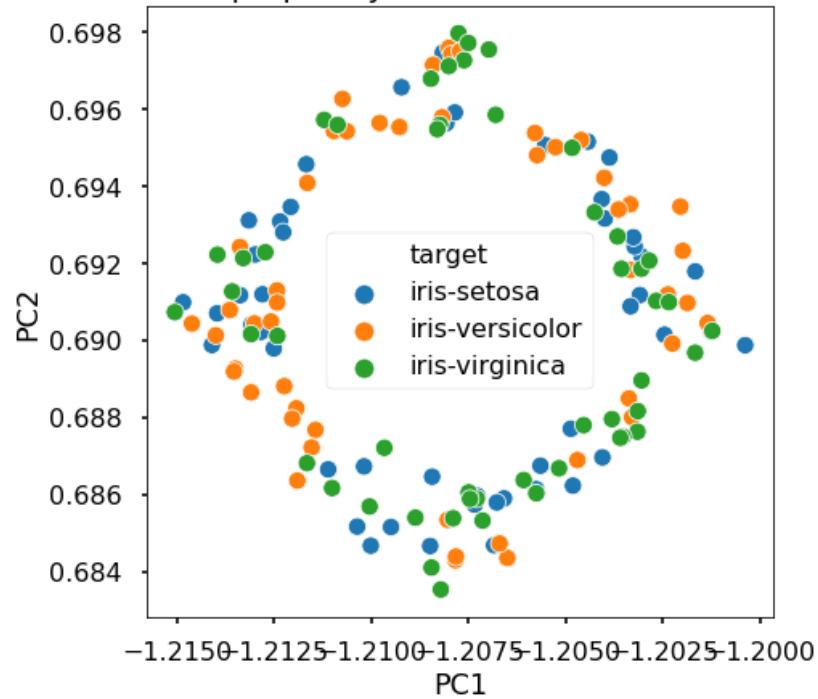
t-SNE visualization with perplexity -- 200, iterations -- 3500 and epsilon -- 200



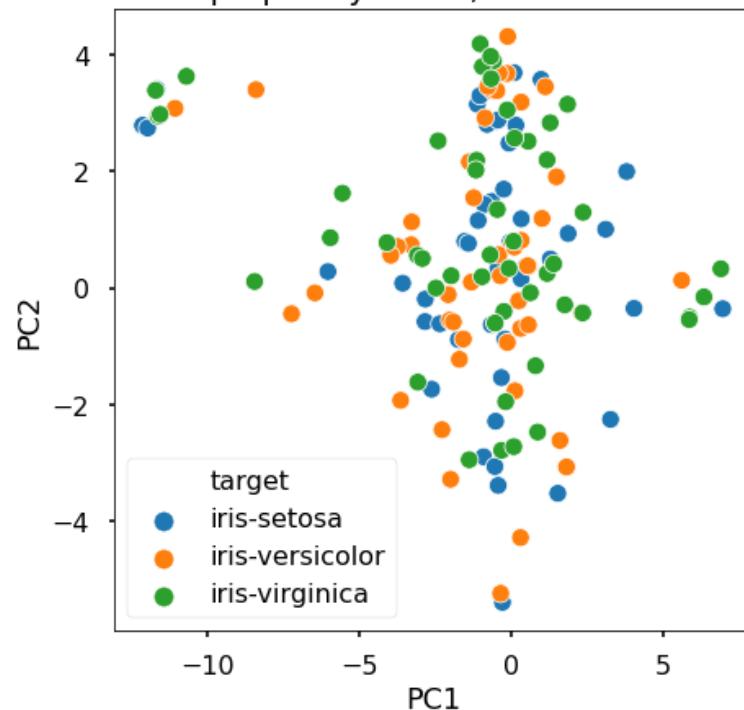
t-SNE visualization with perplexity -- 200, iterations -- 4000 and epsilon -- 200



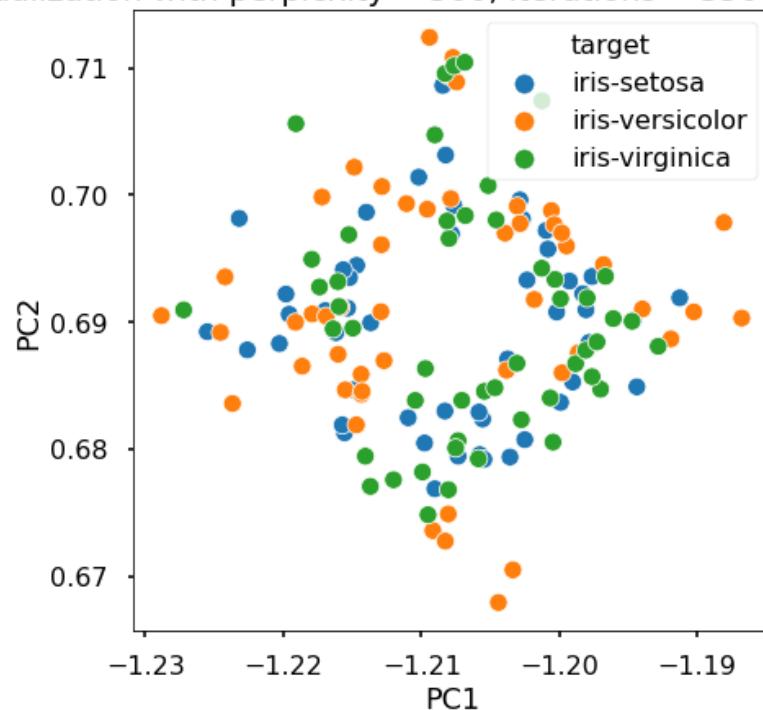
t-SNE visualization with perplexity -- 200, iterations -- 5000 and epsilon -- 200



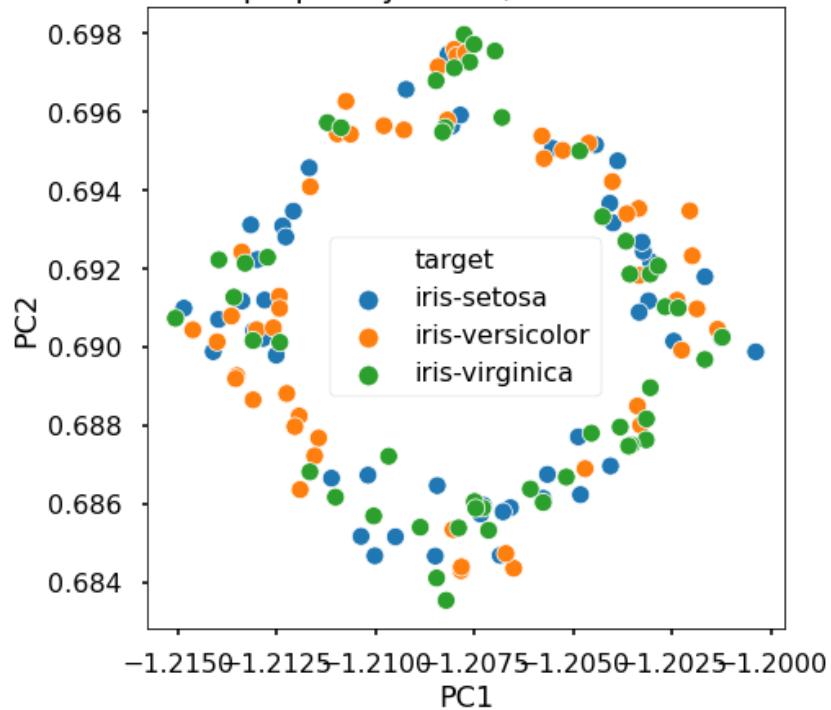
t-SNE visualization with perplexity -- 300, iterations -- 250 and epsilon -- 200



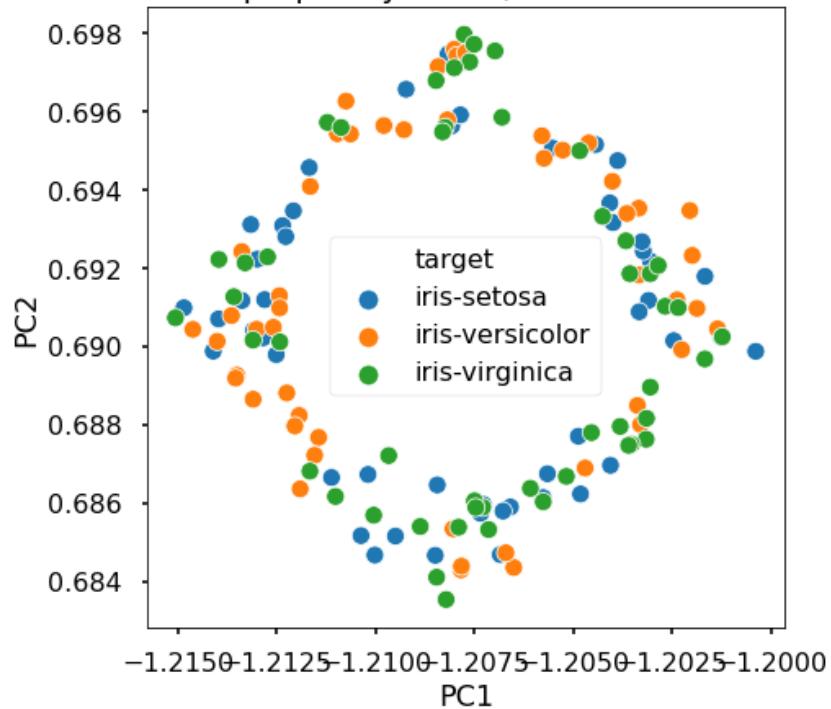
t-SNE visualization with perplexity -- 300, iterations -- 350 and epsilon -- 200



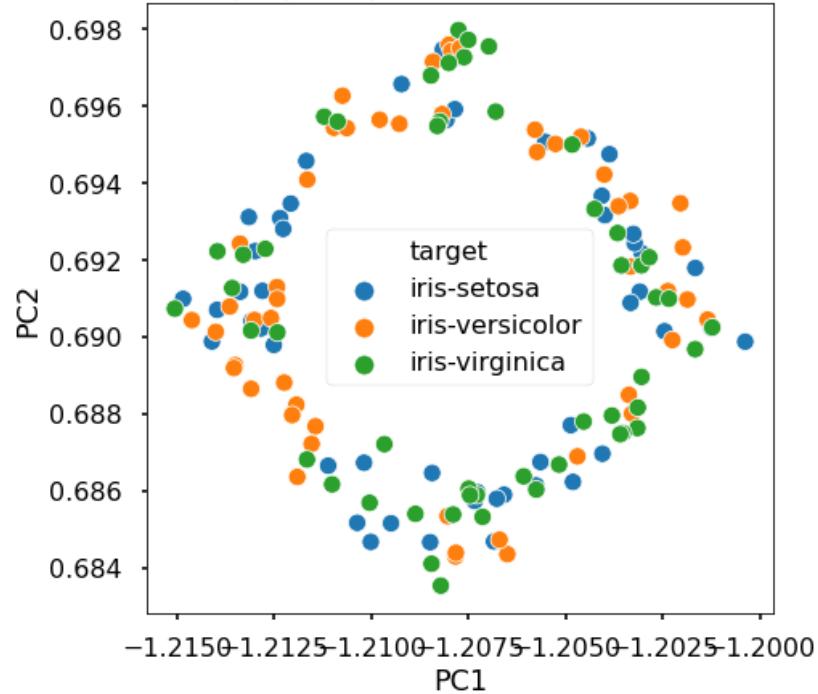
t-SNE visualization with perplexity -- 300, iterations -- 500 and epsilon -- 200



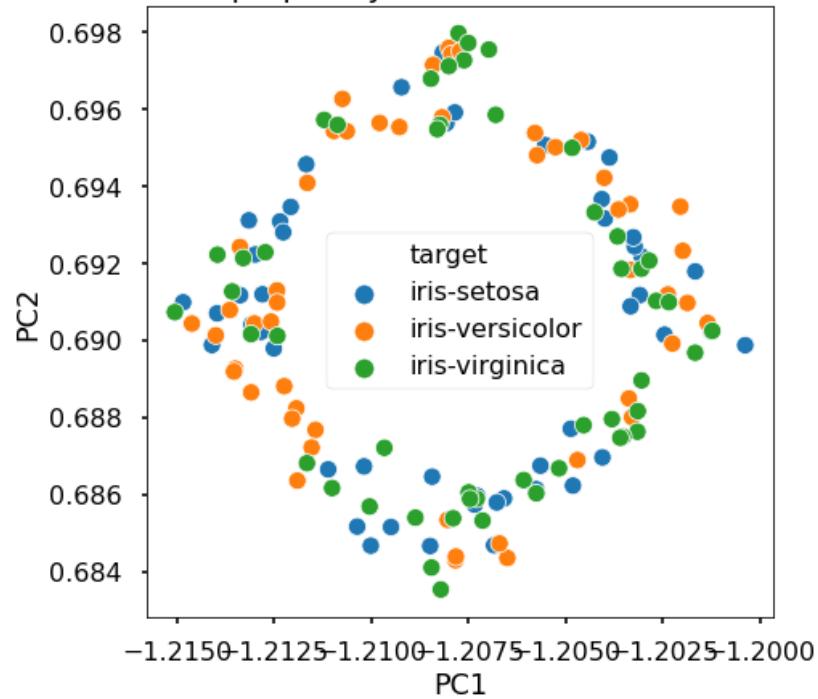
t-SNE visualization with perplexity -- 300, iterations -- 750 and epsilon -- 200



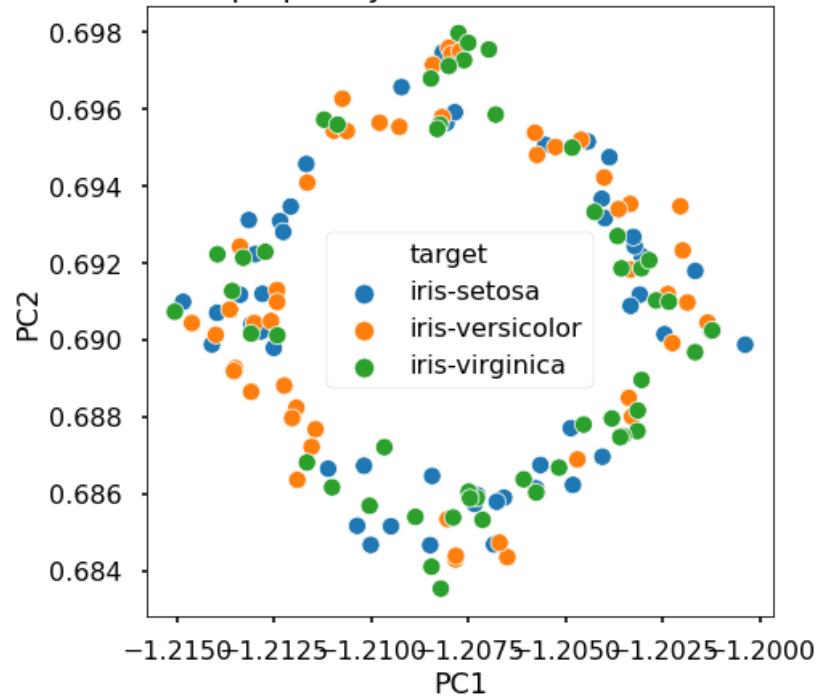
t-SNE visualization with perplexity -- 300, iterations -- 1000 and epsilon -- 200



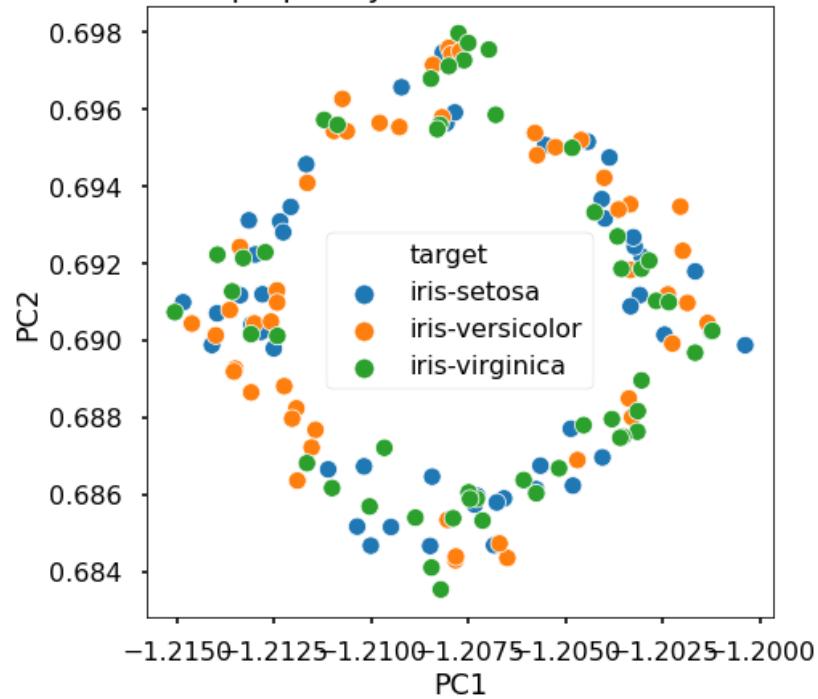
t-SNE visualization with perplexity -- 300, iterations -- 2500 and epsilon -- 200



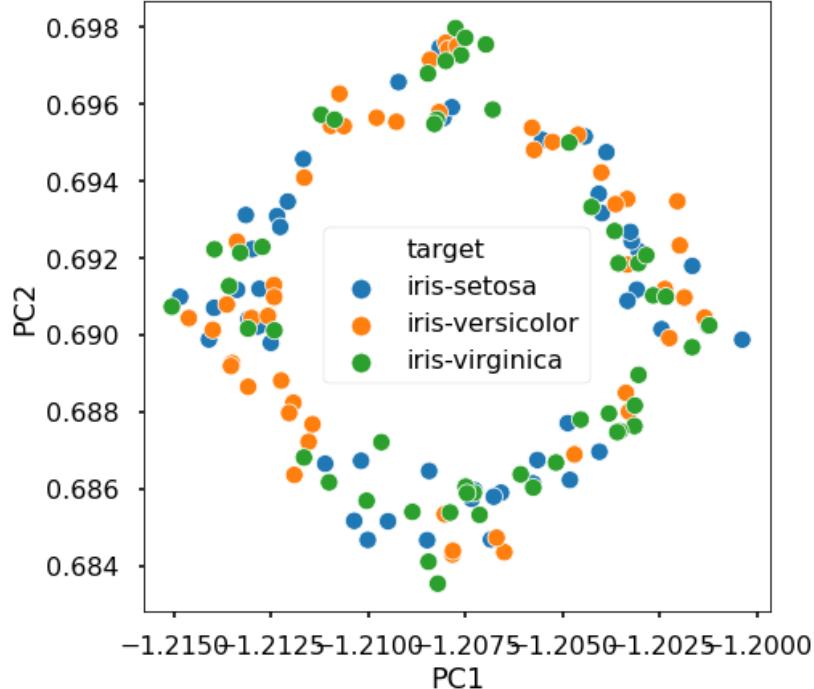
t-SNE visualization with perplexity -- 300, iterations -- 3500 and epsilon -- 200



t-SNE visualization with perplexity -- 300, iterations -- 4000 and epsilon -- 200



t-SNE visualization with perplexity -- 300, iterations -- 5000 and epsilon -- 200



IRIS:CASE-II-Multiple_runs

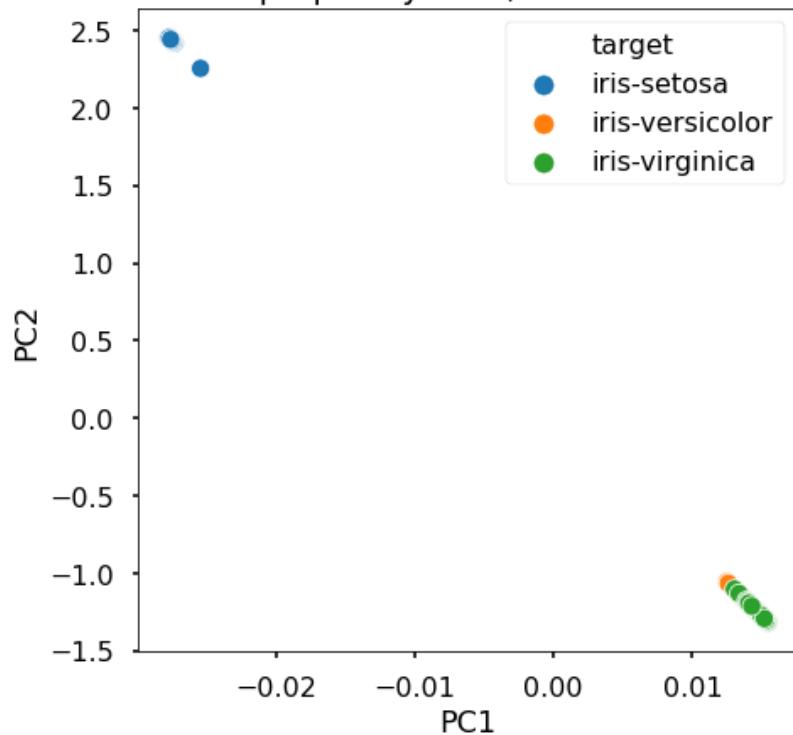
Running t-SNE on a range of Learning rates(epsilon) and iterations with fixed perplexity and embedding initialization as 'random'

```
In [79]: l_rate = [10,30,50,100,200,400,600]
iterations = [250,350,500,750,1000,2500,3500,4000,5000]

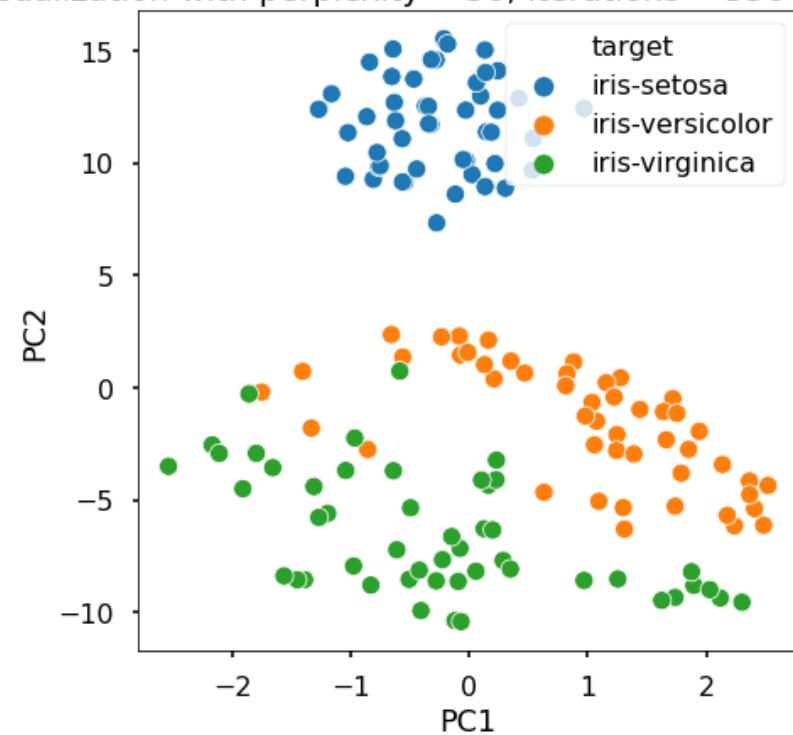
for lidx in range(len(l_rate)):
    for idx in range(len(iterations)):
        tsne_iris_1 = TSNE(n_components=2,perplexity=30,learning_rate=l_rate[lidx],n_init='random',n_jobs=-1,random_state=23)
        iris1_tsne_pcmps = pd.DataFrame(tsne_iris_1.fit_transform(iris_stand_df.iloc))
        iris1_tsne_pcmps = pd.concat([iris1_tsne_pcmps,iris_stand_df['target']],axis=1)
        with plt.style.context('seaborn-poster'):
            plt.figure(figsize=(7,7))
            sns.scatterplot(data=iris1_tsne_pcmps,x='PC1',y='PC2',hue='target')
            plt.title("t-SNE visualization with perplexity -- {0}, iterations -- {1}")

plt.show()
```

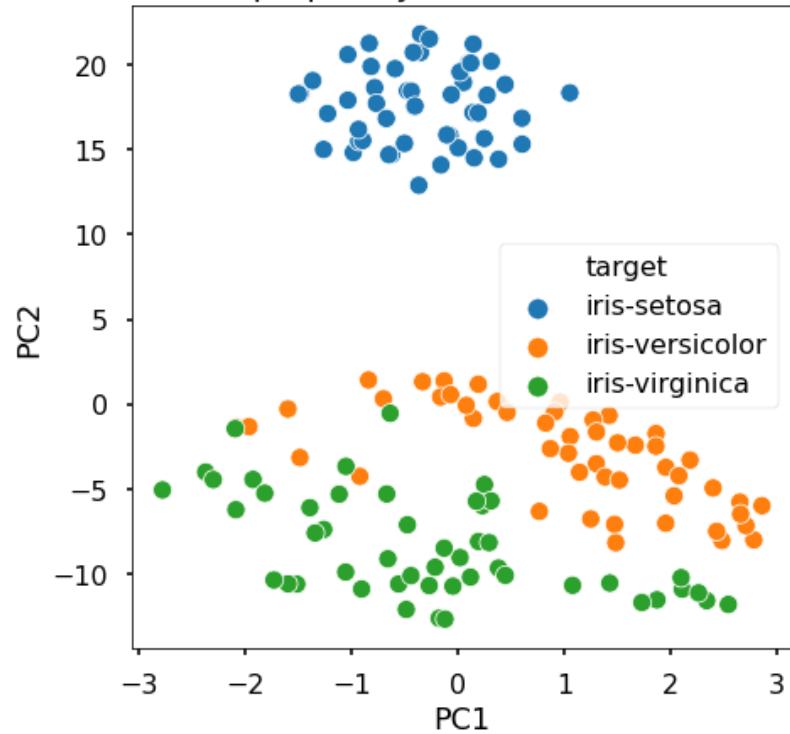
t-SNE visualization with perplexity -- 30, iterations -- 250 and epsilon -- 10



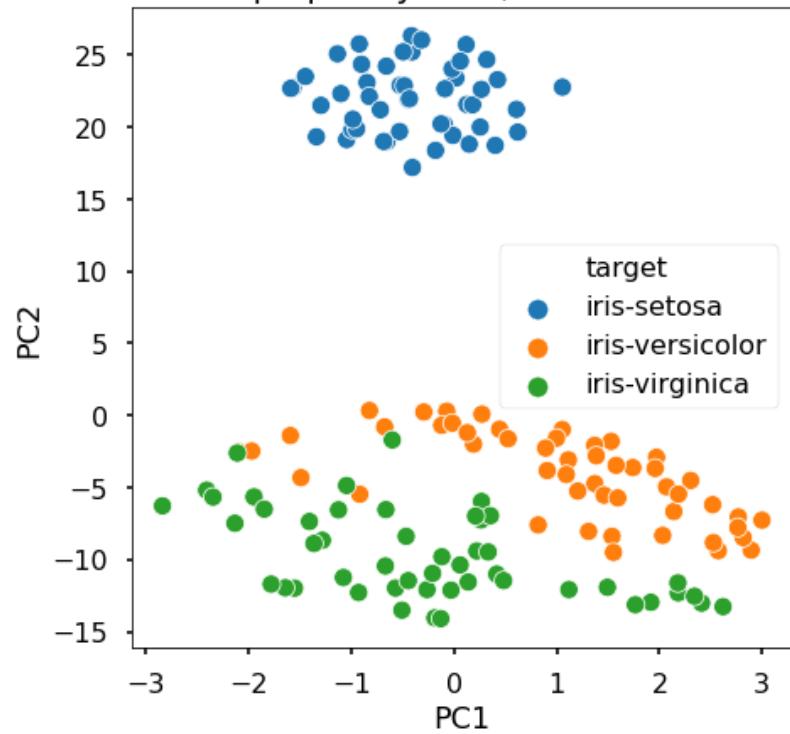
t-SNE visualization with perplexity -- 30, iterations -- 350 and epsilon -- 10



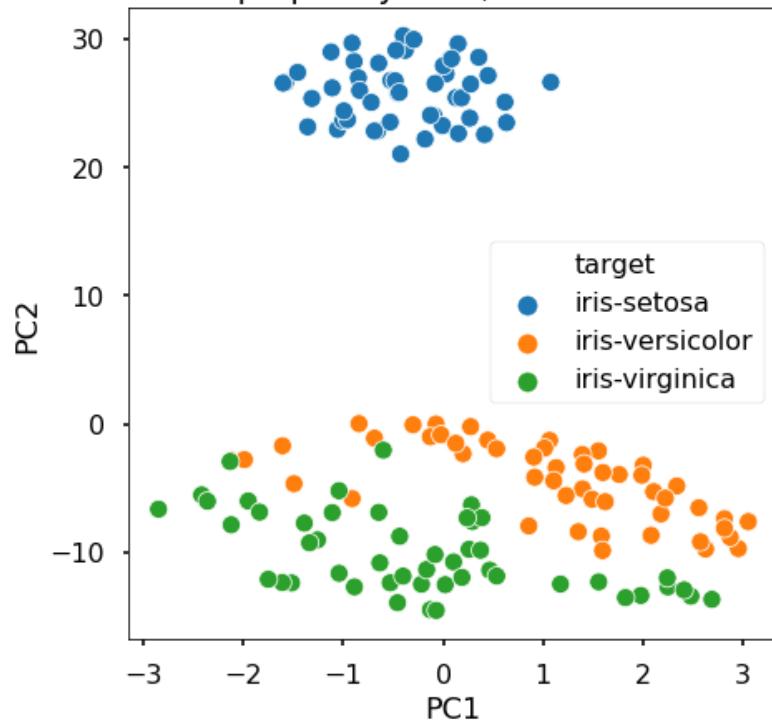
t-SNE visualization with perplexity -- 30, iterations -- 500 and epsilon -- 10



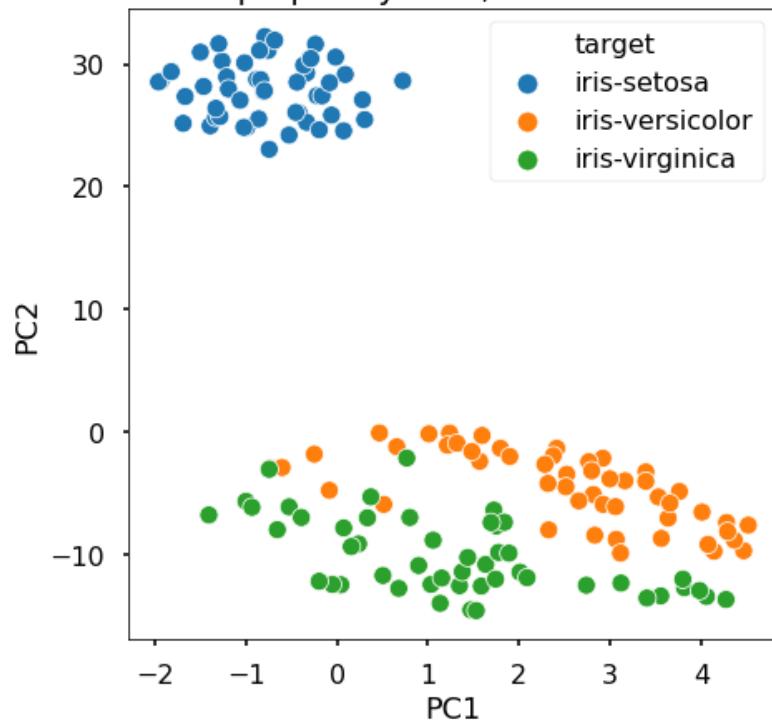
t-SNE visualization with perplexity -- 30, iterations -- 750 and epsilon -- 10



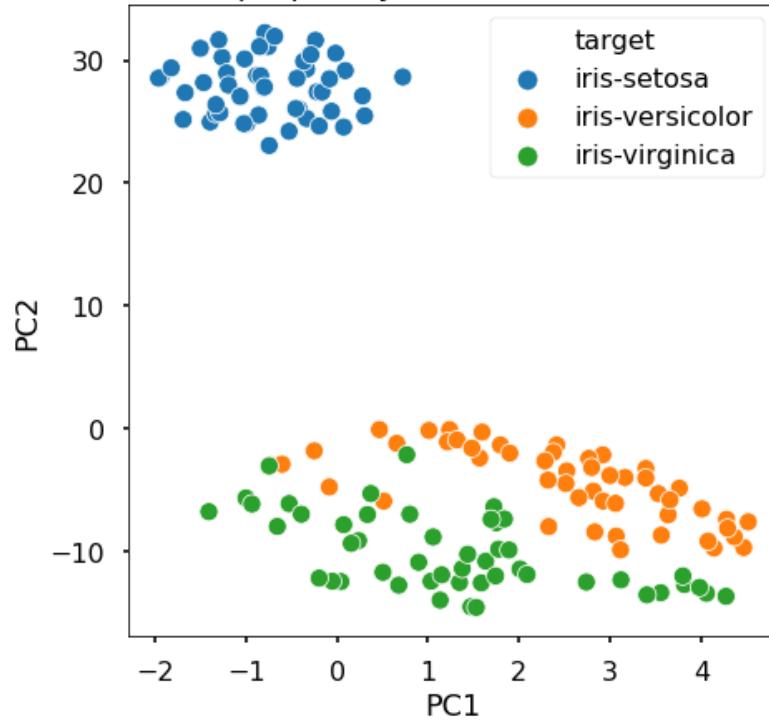
t-SNE visualization with perplexity -- 30, iterations -- 1000 and epsilon -- 10



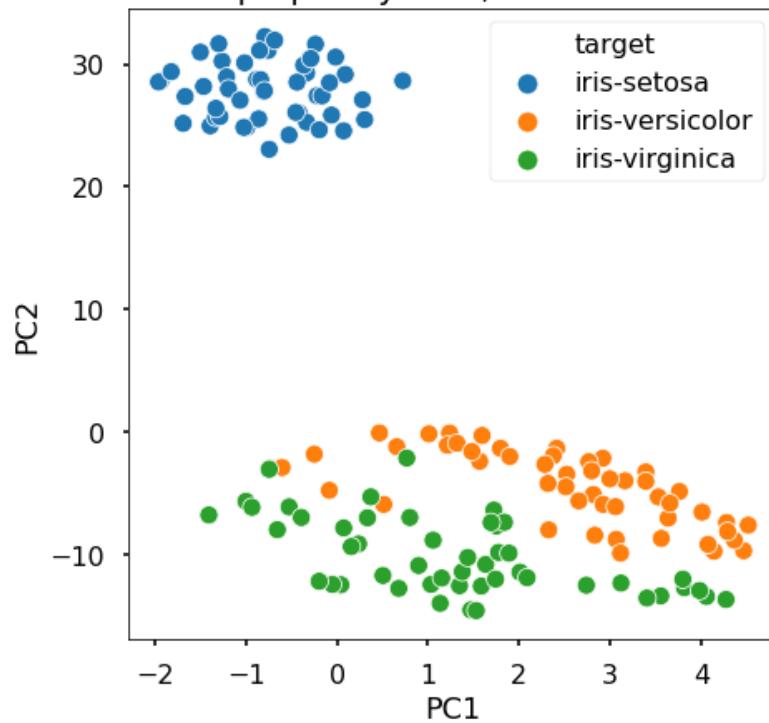
t-SNE visualization with perplexity -- 30, iterations -- 2500 and epsilon -- 10



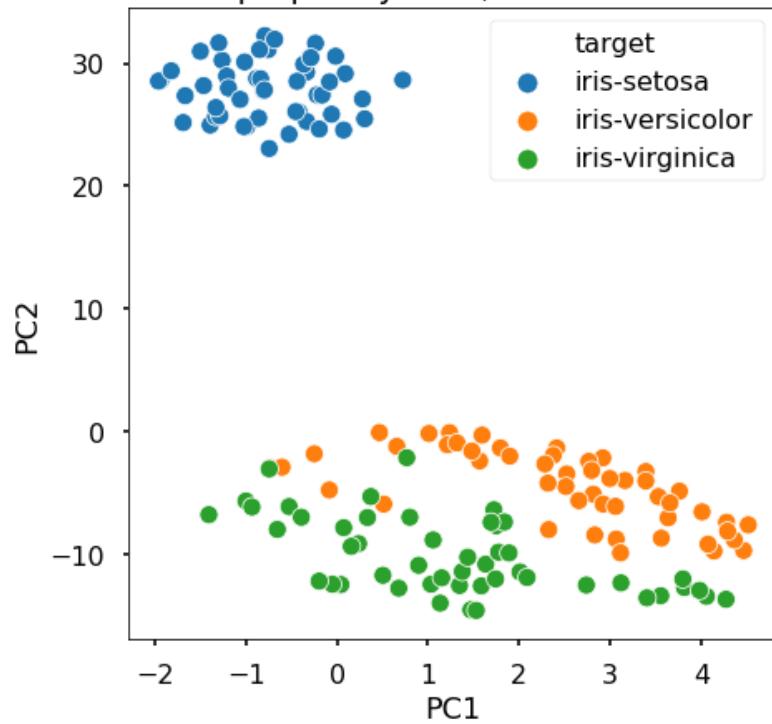
t-SNE visualization with perplexity -- 30, iterations -- 3500 and epsilon -- 10



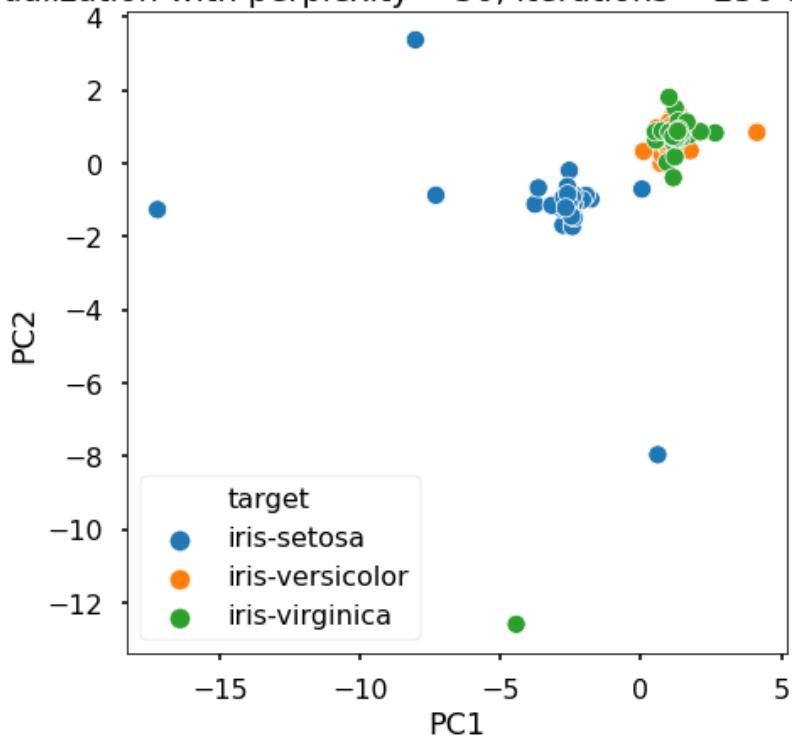
t-SNE visualization with perplexity -- 30, iterations -- 4000 and epsilon -- 10



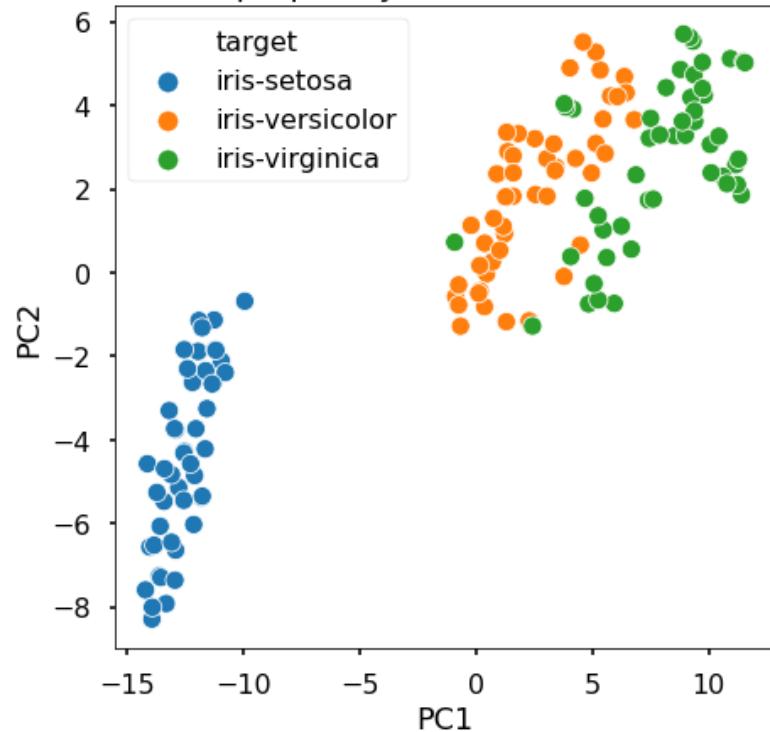
t-SNE visualization with perplexity -- 30, iterations -- 5000 and epsilon -- 10



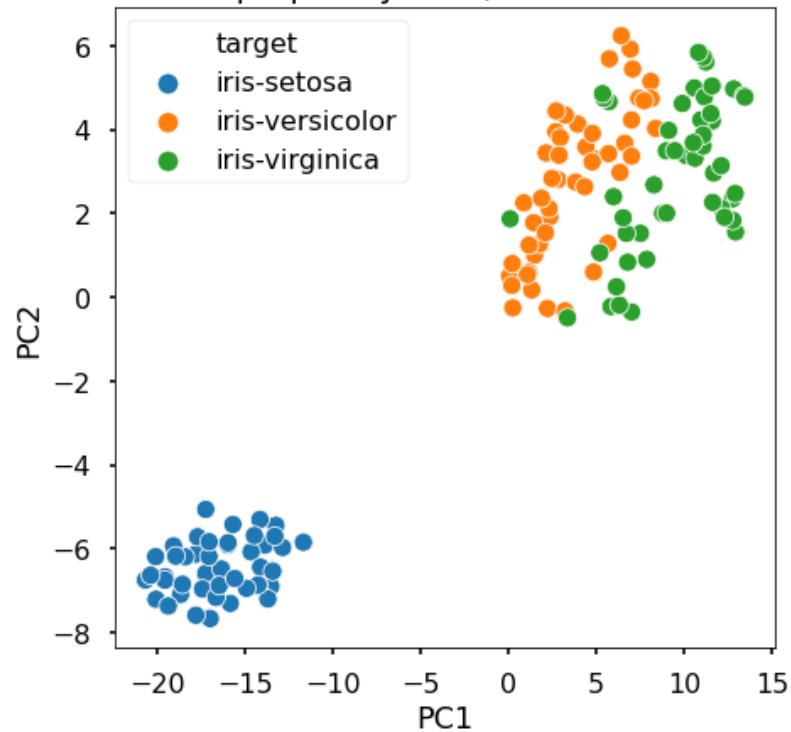
t-SNE visualization with perplexity -- 30, iterations -- 250 and epsilon -- 50



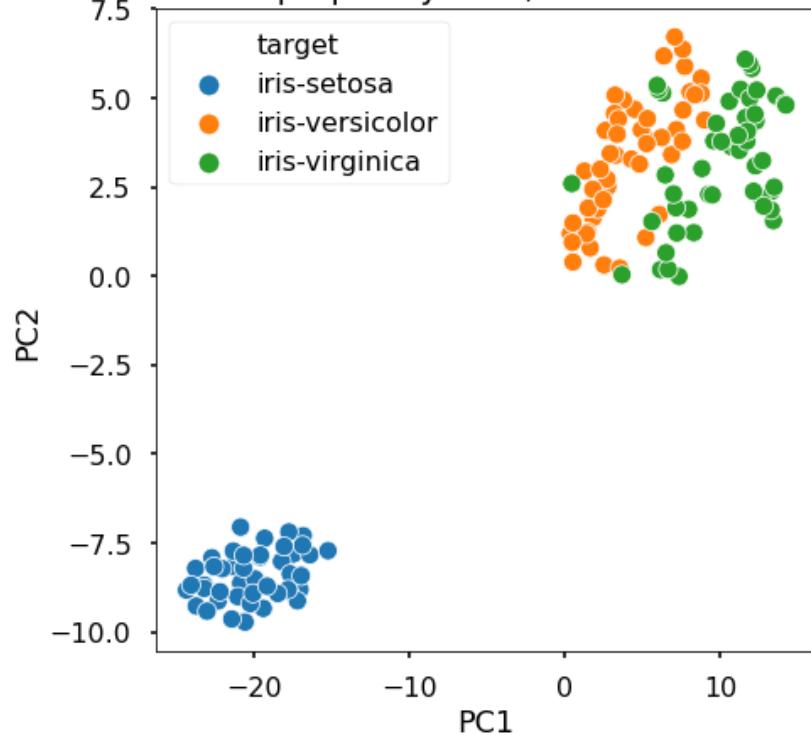
t-SNE visualization with perplexity -- 30, iterations -- 350 and epsilon -- 50



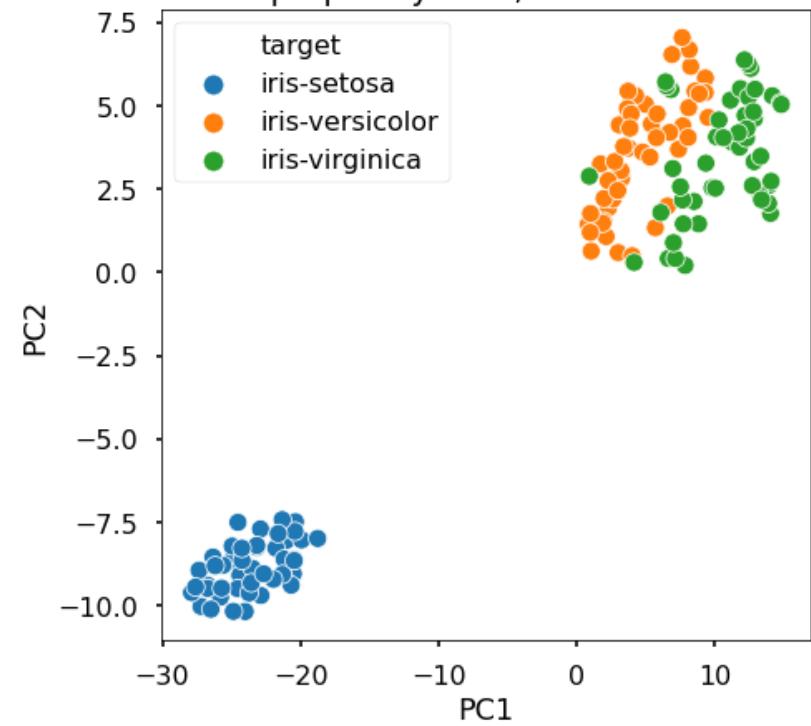
t-SNE visualization with perplexity -- 30, iterations -- 500 and epsilon -- 50



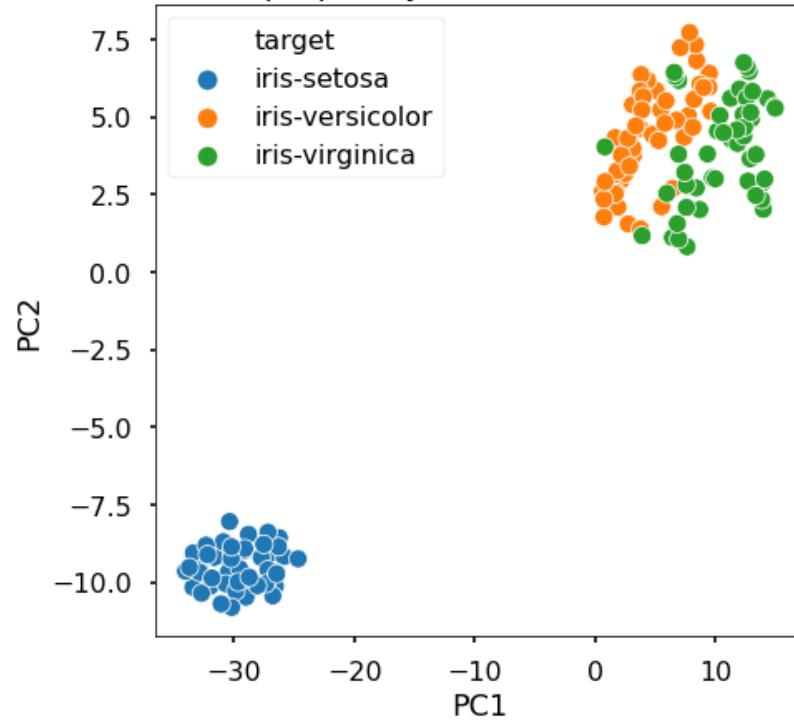
t-SNE visualization with perplexity -- 30, iterations -- 750 and epsilon -- 50



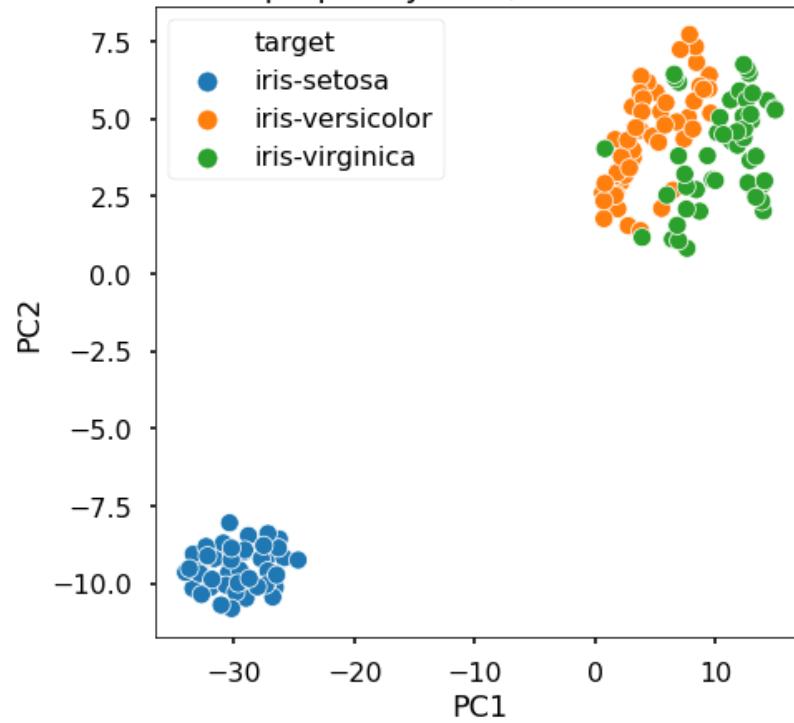
t-SNE visualization with perplexity -- 30, iterations -- 1000 and epsilon -- 50



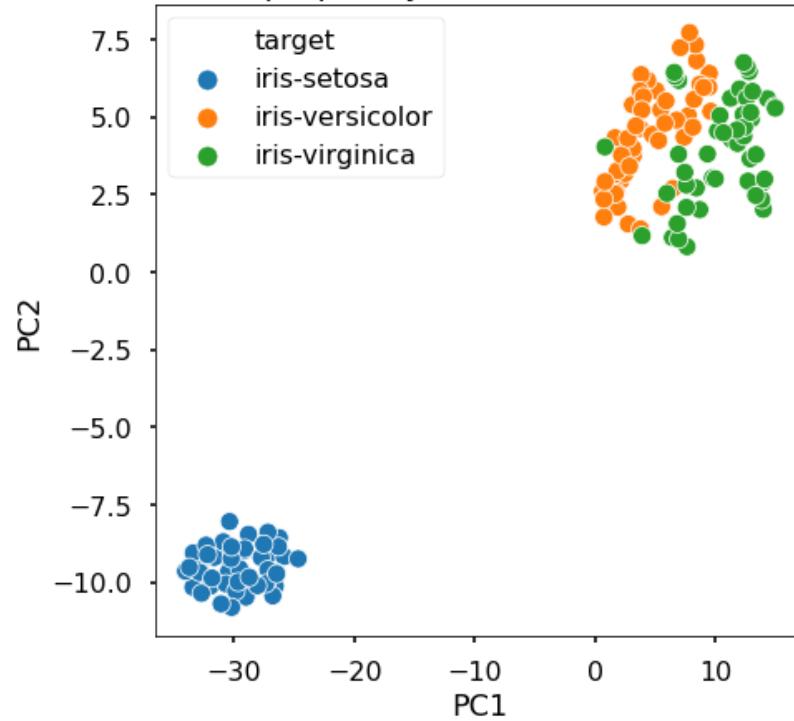
t-SNE visualization with perplexity -- 30, iterations -- 2500 and epsilon -- 50



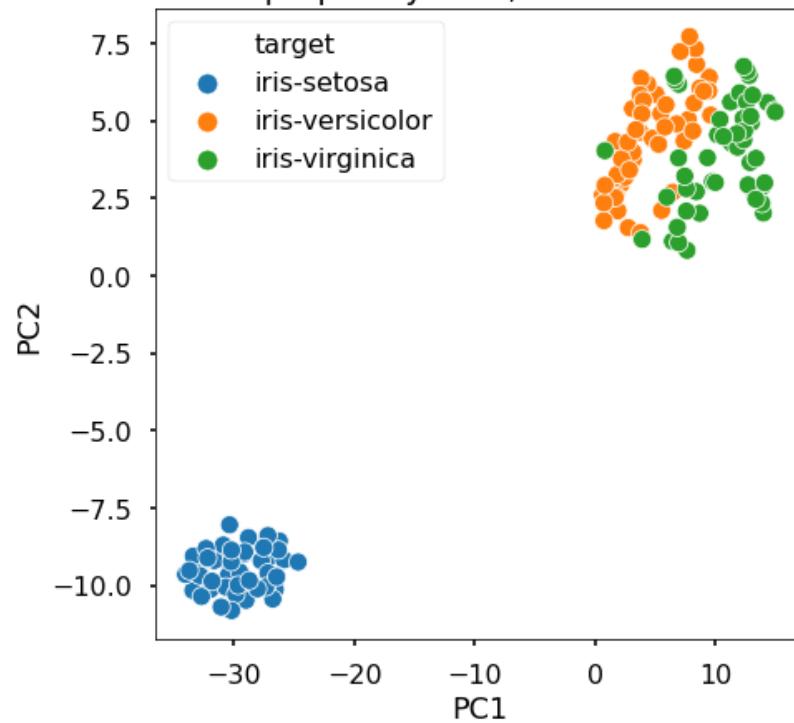
t-SNE visualization with perplexity -- 30, iterations -- 3500 and epsilon -- 50



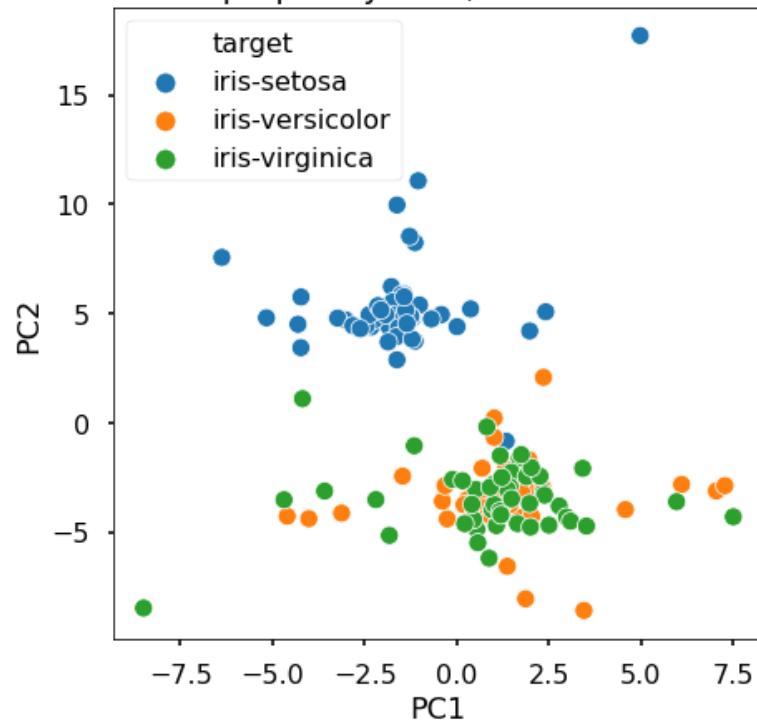
t-SNE visualization with perplexity -- 30, iterations -- 4000 and epsilon -- 50



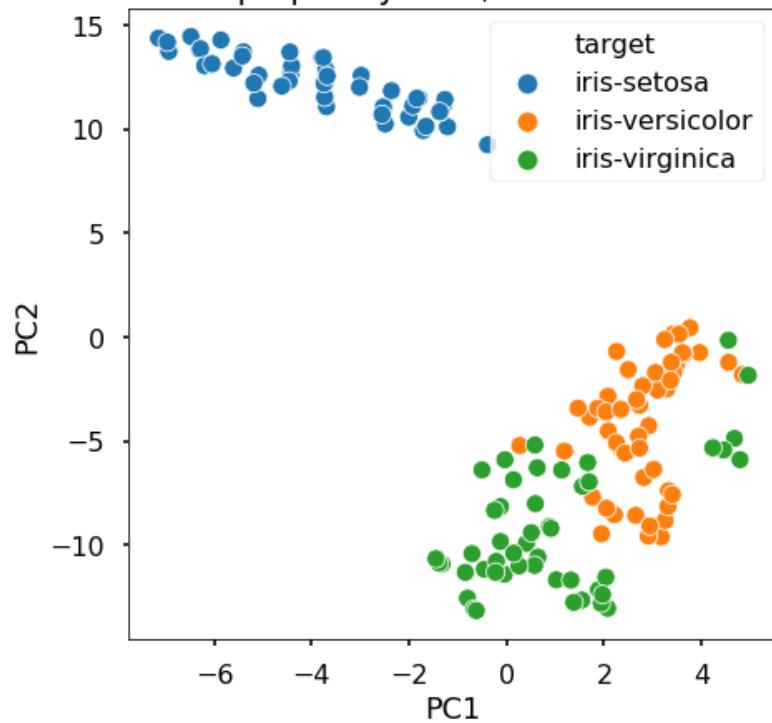
t-SNE visualization with perplexity -- 30, iterations -- 5000 and epsilon -- 50



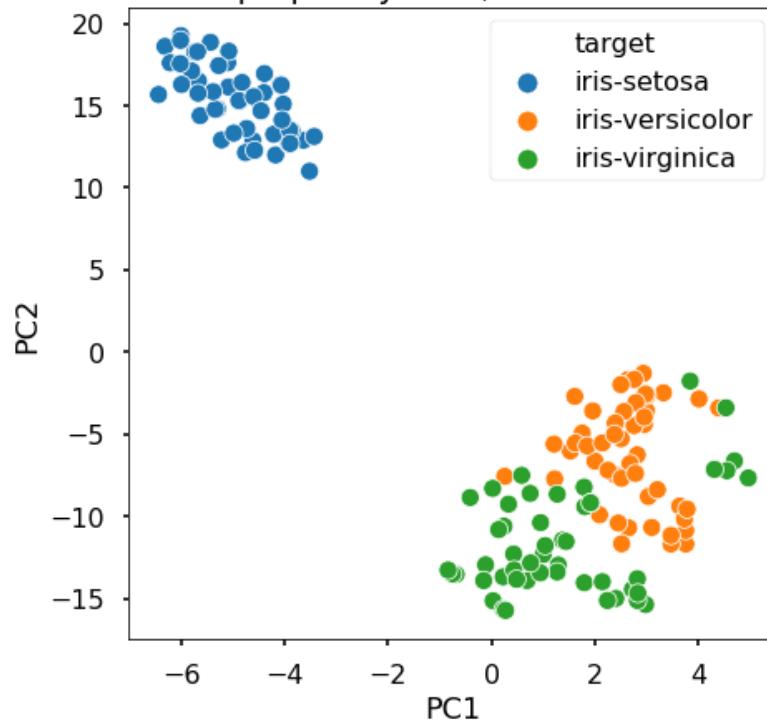
t-SNE visualization with perplexity -- 30, iterations -- 250 and epsilon -- 100



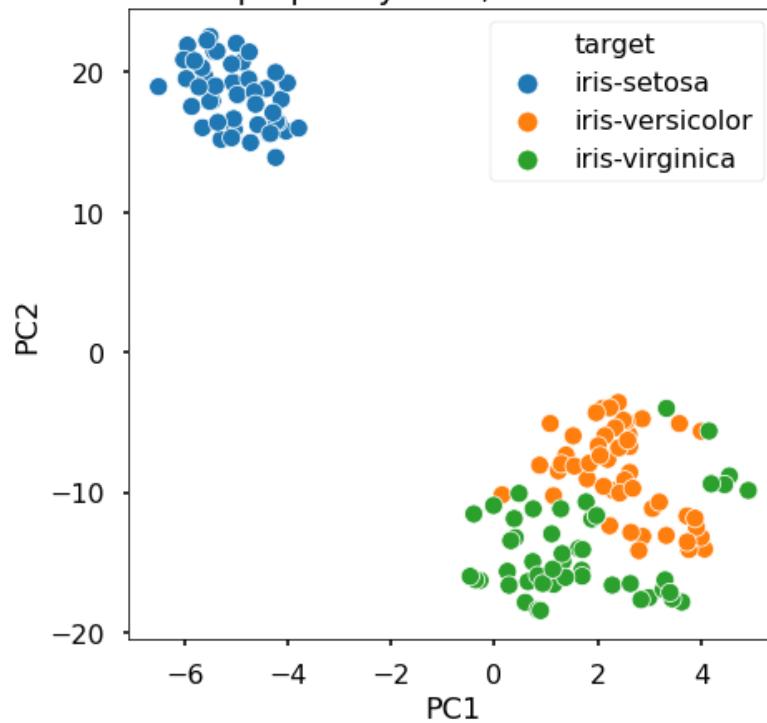
t-SNE visualization with perplexity -- 30, iterations -- 350 and epsilon -- 100



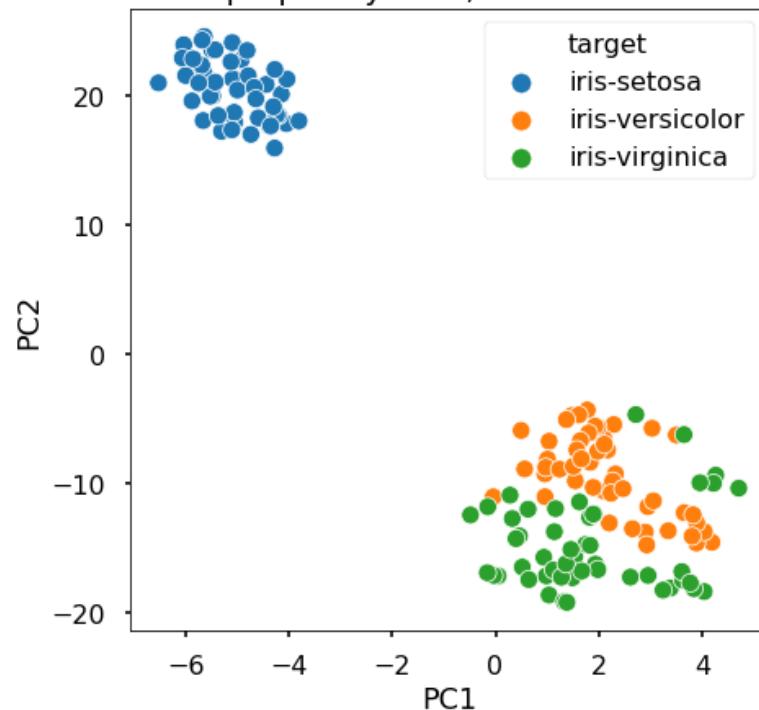
t-SNE visualization with perplexity -- 30, iterations -- 500 and epsilon -- 100



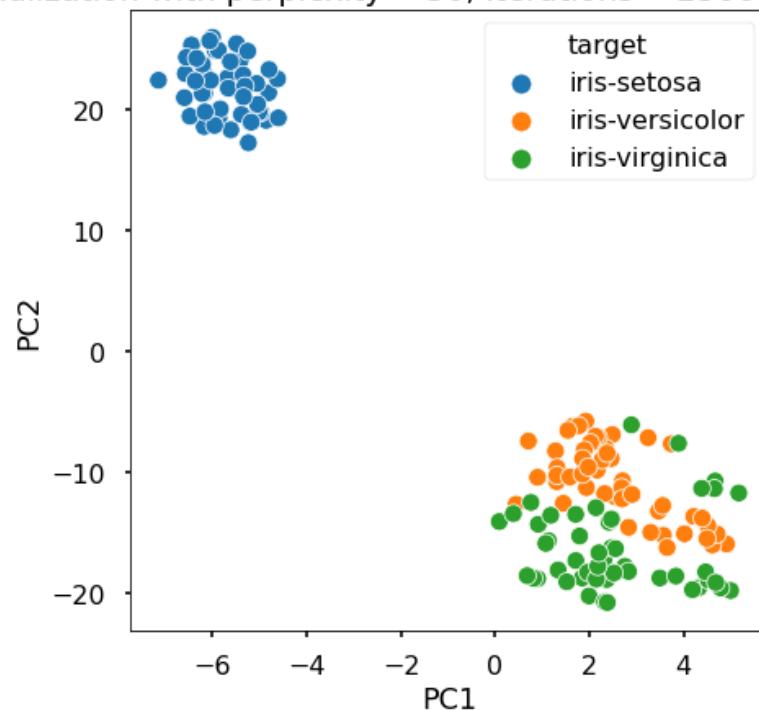
t-SNE visualization with perplexity -- 30, iterations -- 750 and epsilon -- 100



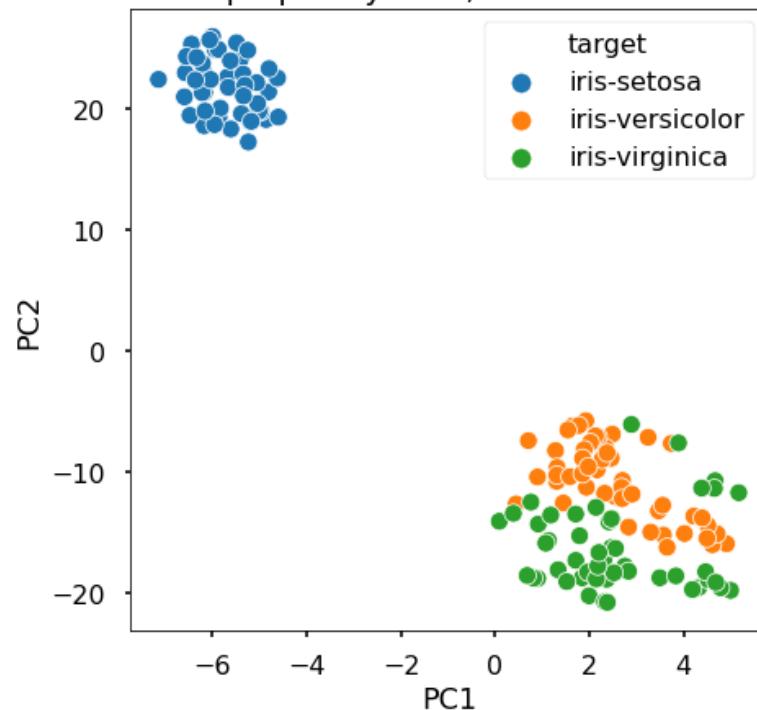
t-SNE visualization with perplexity -- 30, iterations -- 1000 and epsilon -- 100



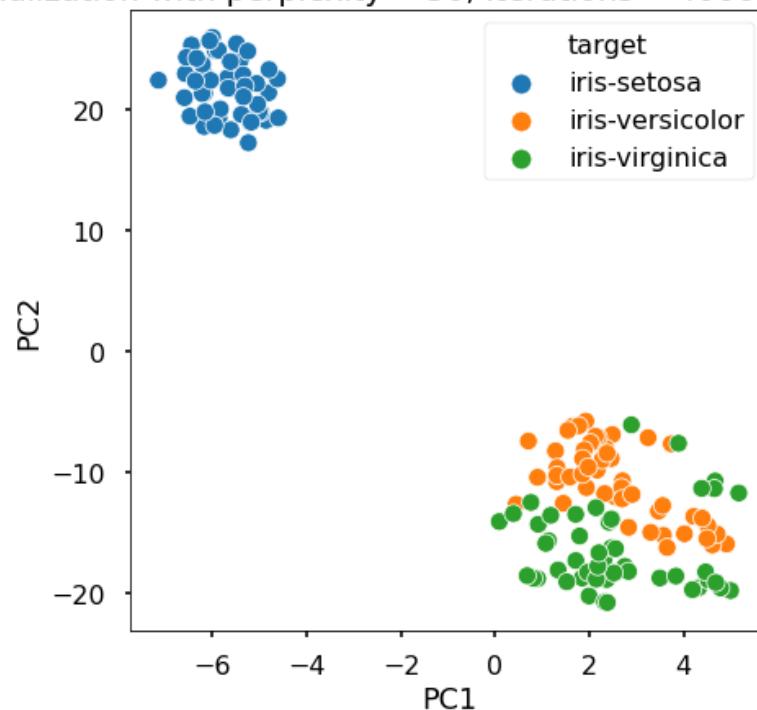
t-SNE visualization with perplexity -- 30, iterations -- 2500 and epsilon -- 100



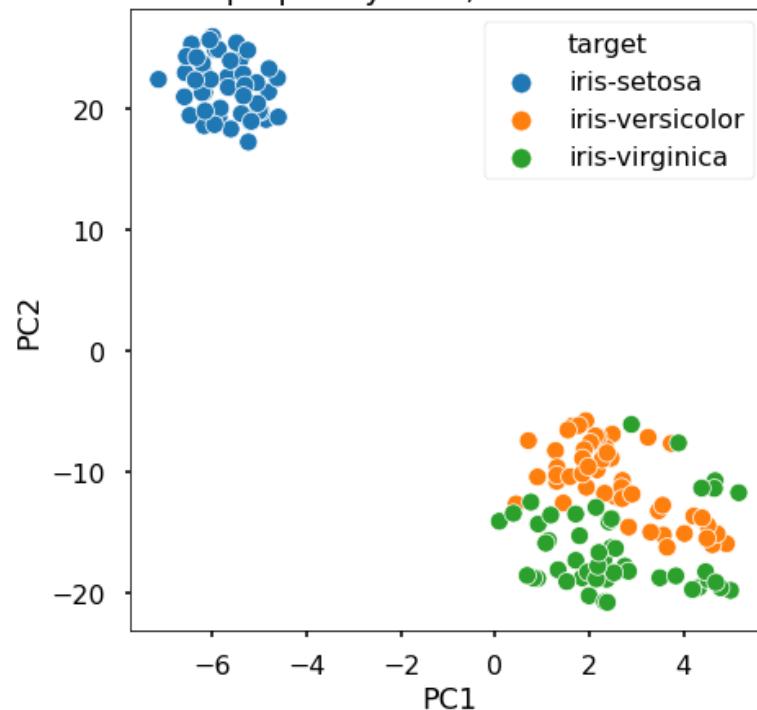
t-SNE visualization with perplexity -- 30, iterations -- 3500 and epsilon -- 100



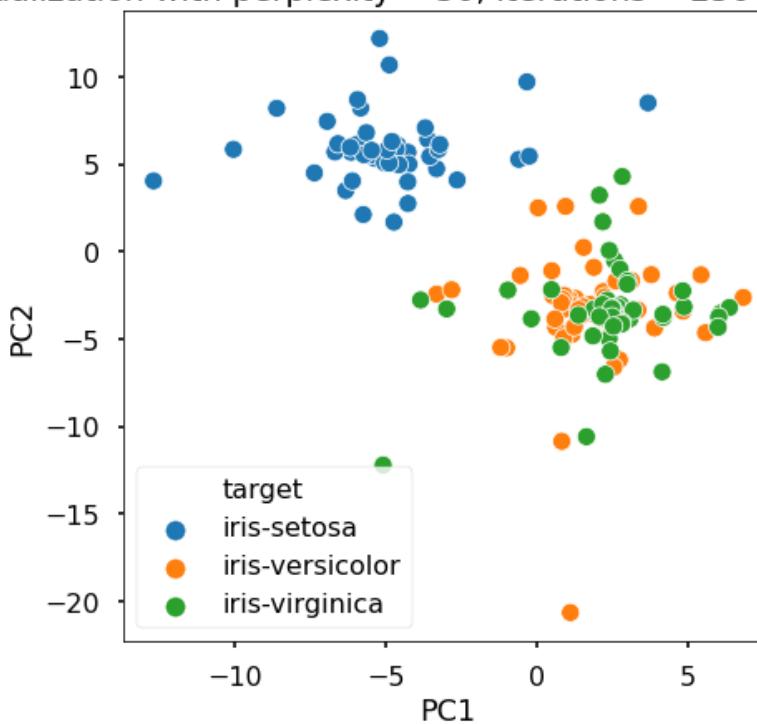
t-SNE visualization with perplexity -- 30, iterations -- 4000 and epsilon -- 100



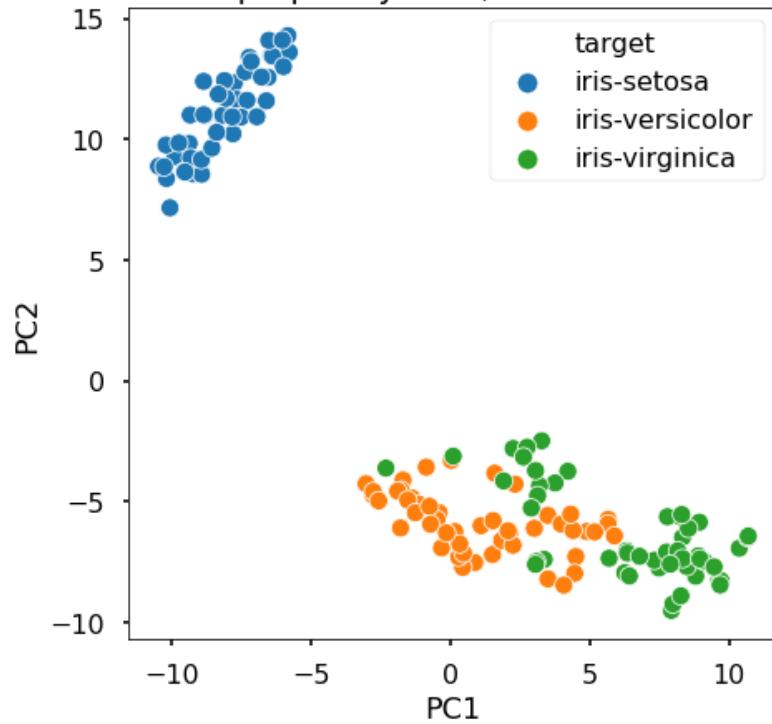
t-SNE visualization with perplexity -- 30, iterations -- 5000 and epsilon -- 100



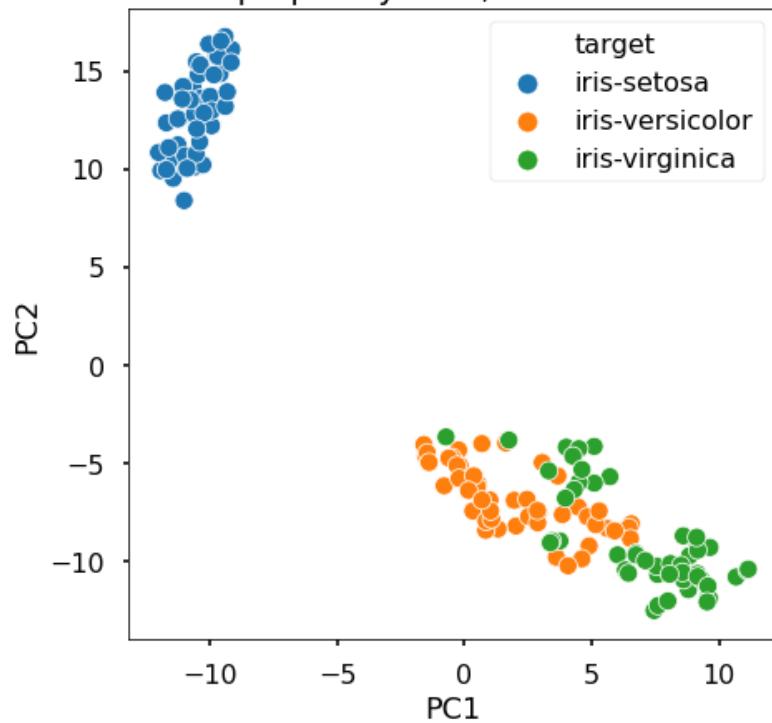
t-SNE visualization with perplexity -- 30, iterations -- 250 and epsilon -- 150



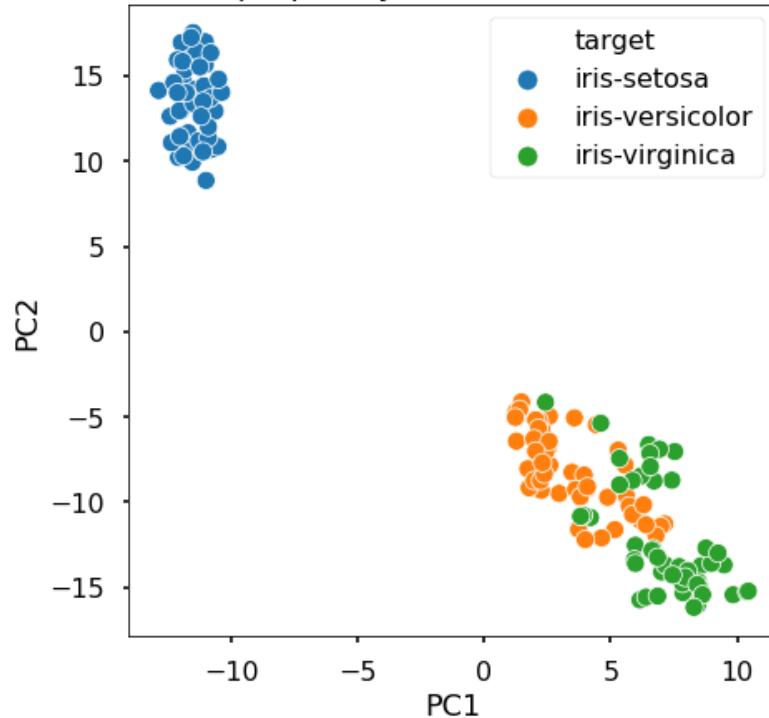
t-SNE visualization with perplexity -- 30, iterations -- 350 and epsilon -- 150



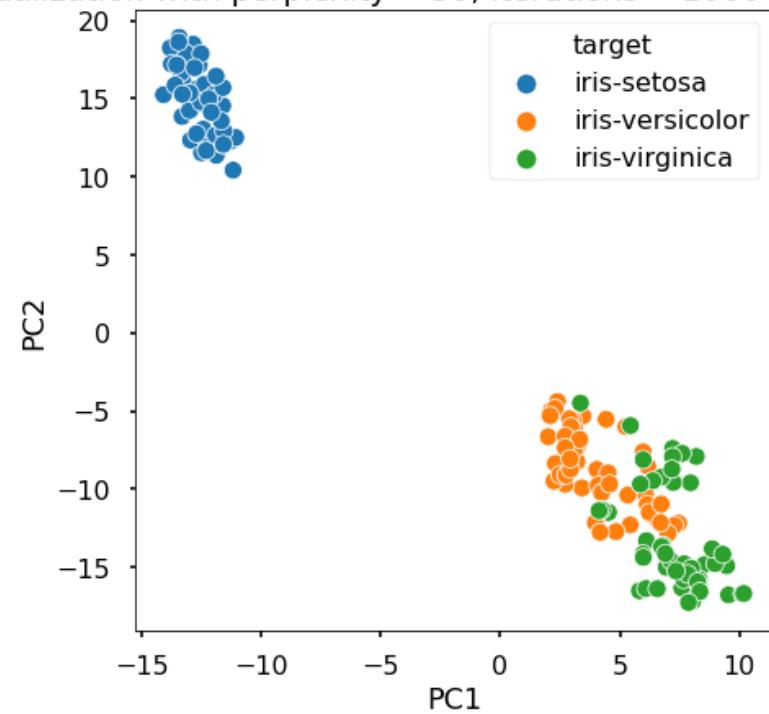
t-SNE visualization with perplexity -- 30, iterations -- 500 and epsilon -- 150



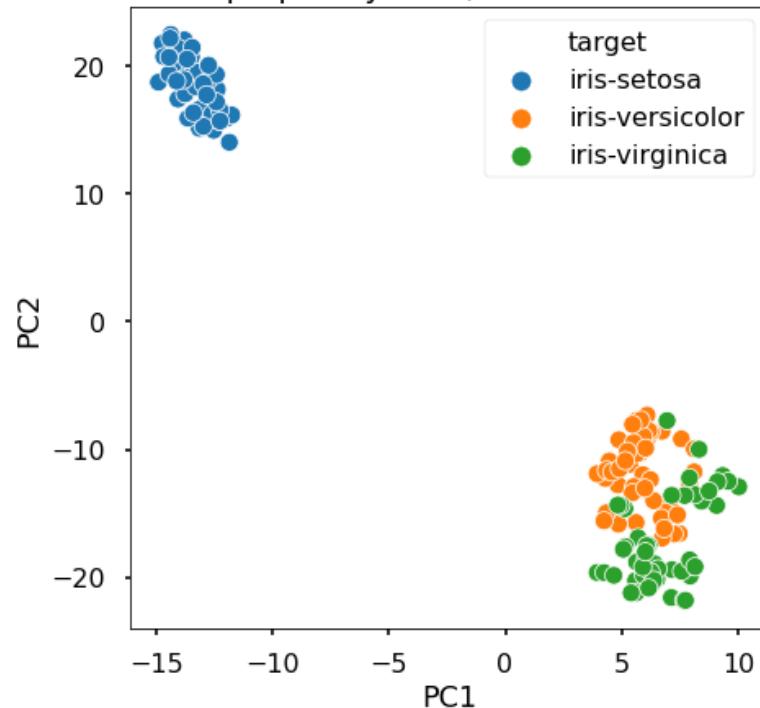
t-SNE visualization with perplexity -- 30, iterations -- 750 and epsilon -- 150



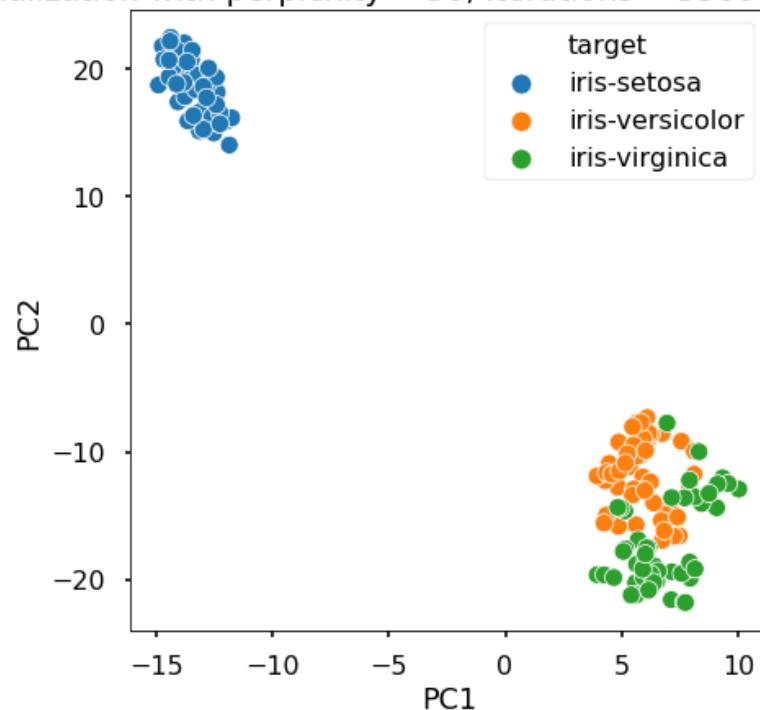
t-SNE visualization with perplexity -- 30, iterations -- 1000 and epsilon -- 150



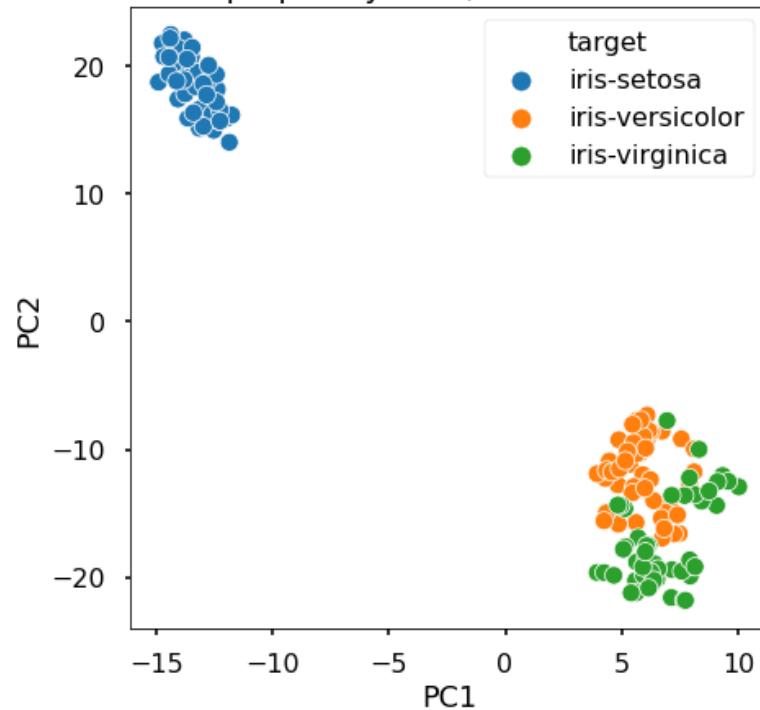
t-SNE visualization with perplexity -- 30, iterations -- 2500 and epsilon -- 150



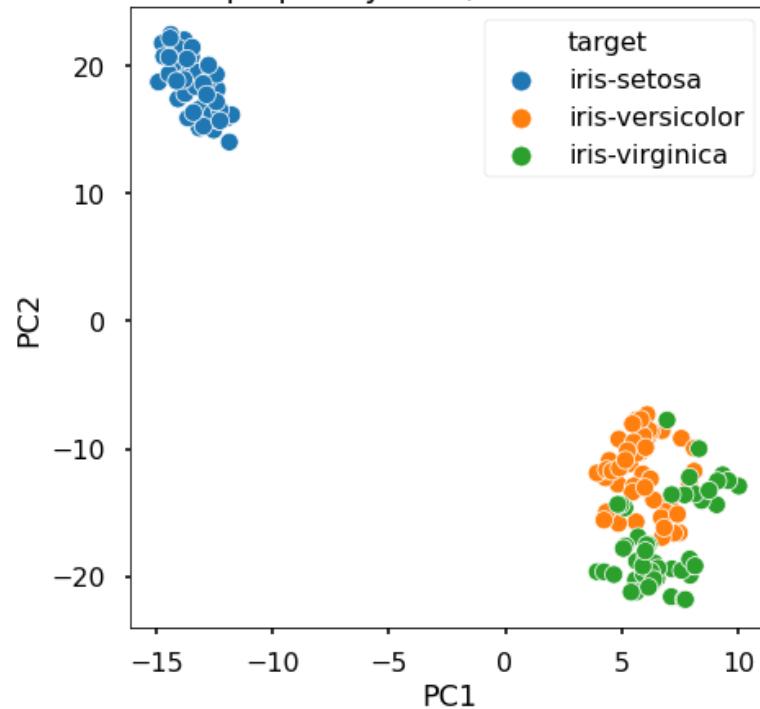
t-SNE visualization with perplexity -- 30, iterations -- 3500 and epsilon -- 150



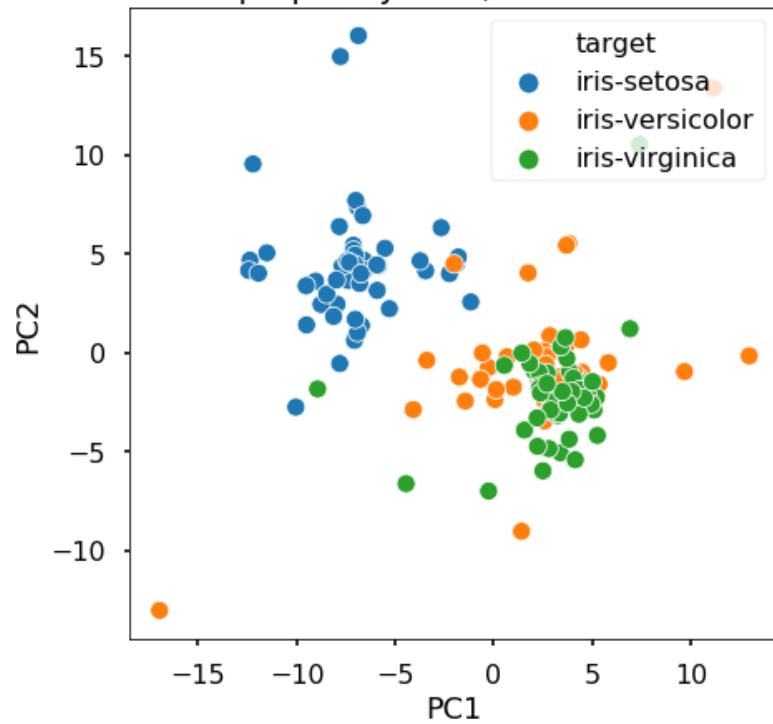
t-SNE visualization with perplexity -- 30, iterations -- 4000 and epsilon -- 150



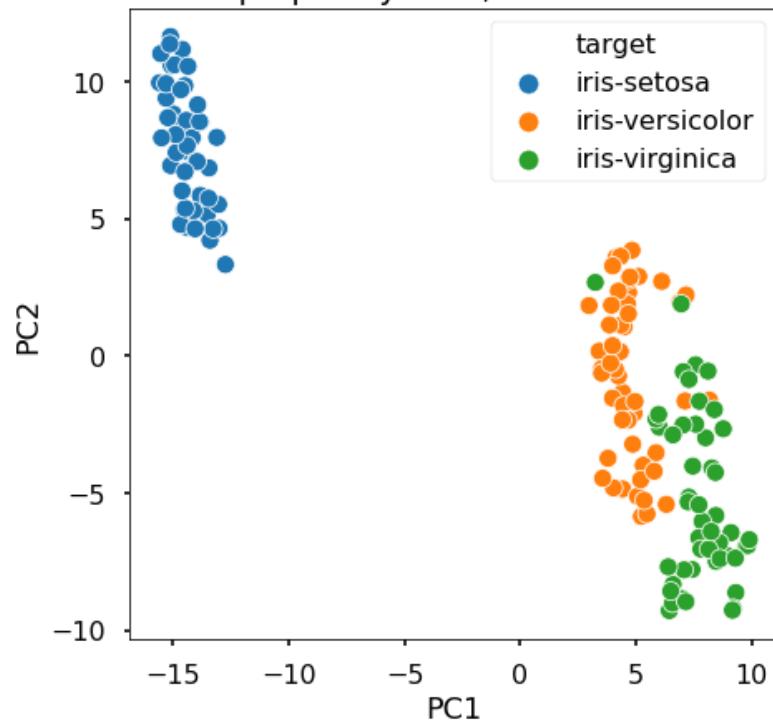
t-SNE visualization with perplexity -- 30, iterations -- 5000 and epsilon -- 150



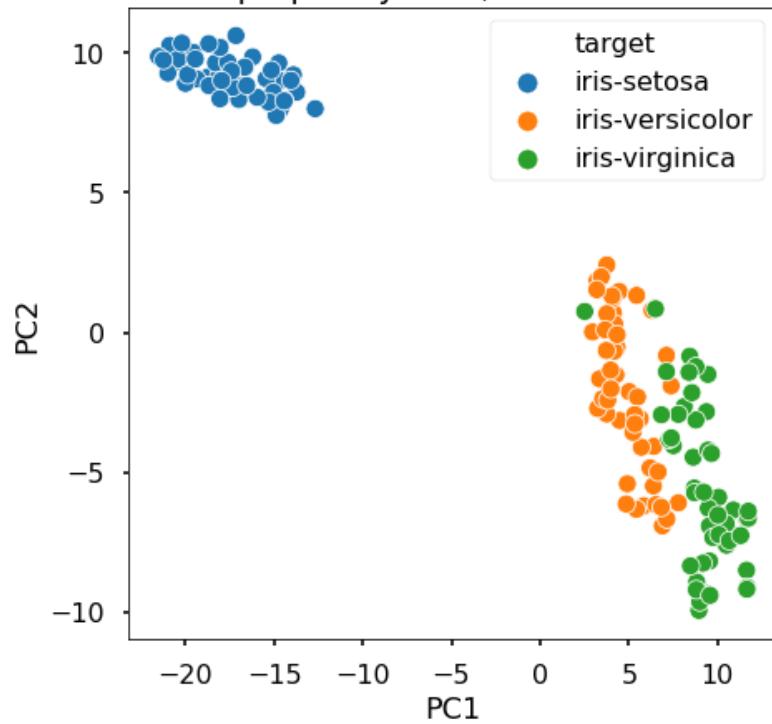
t-SNE visualization with perplexity -- 30, iterations -- 250 and epsilon -- 200



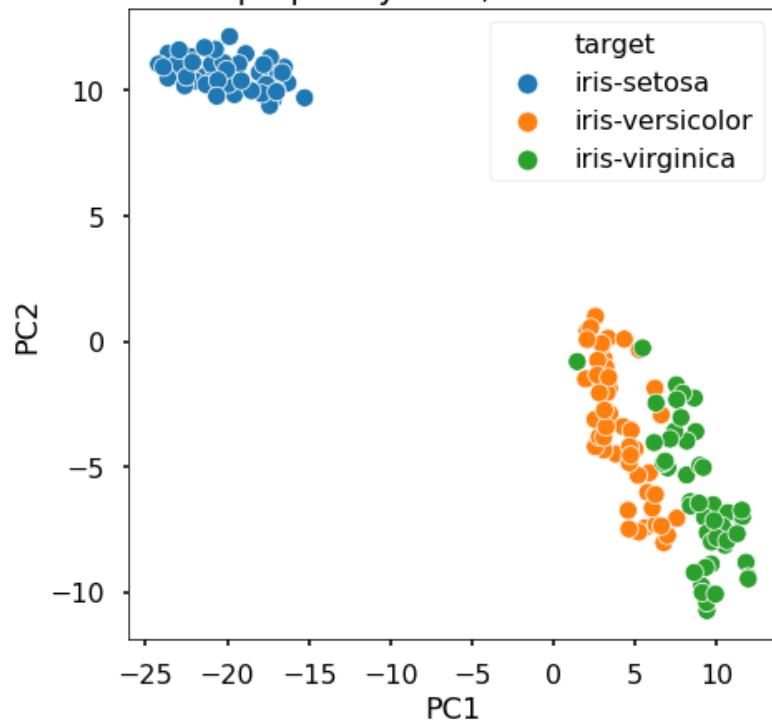
t-SNE visualization with perplexity -- 30, iterations -- 350 and epsilon -- 200



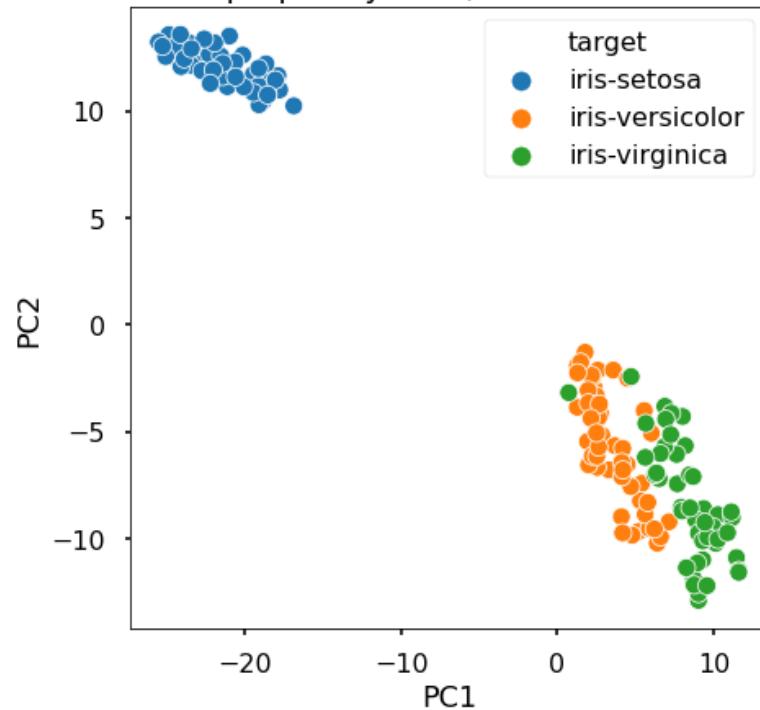
t-SNE visualization with perplexity -- 30, iterations -- 500 and epsilon -- 200



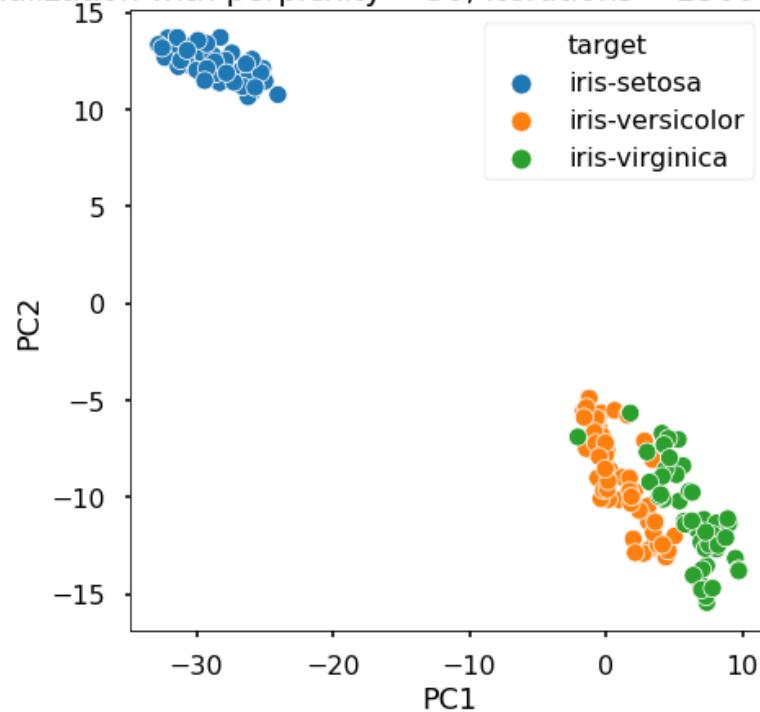
t-SNE visualization with perplexity -- 30, iterations -- 750 and epsilon -- 200



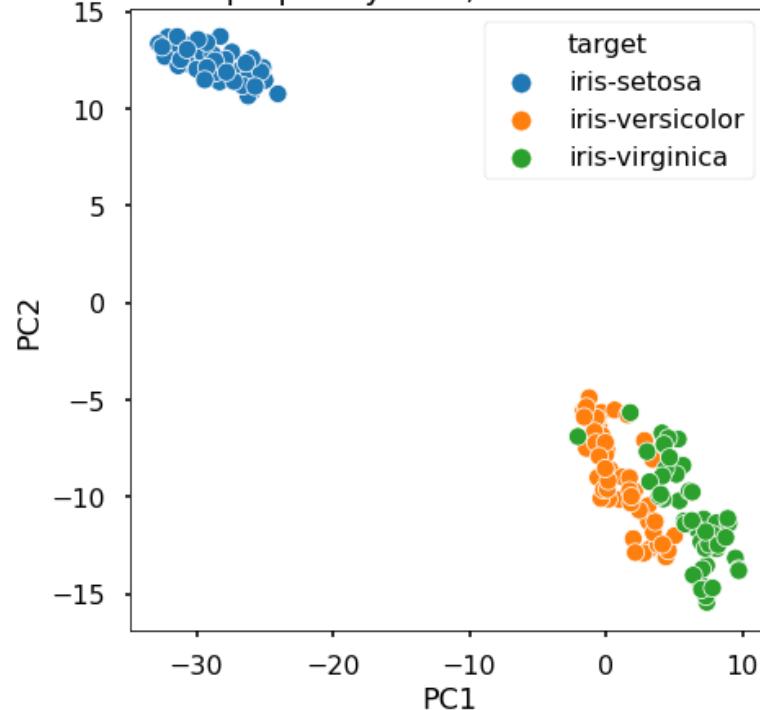
t-SNE visualization with perplexity -- 30, iterations -- 1000 and epsilon -- 200



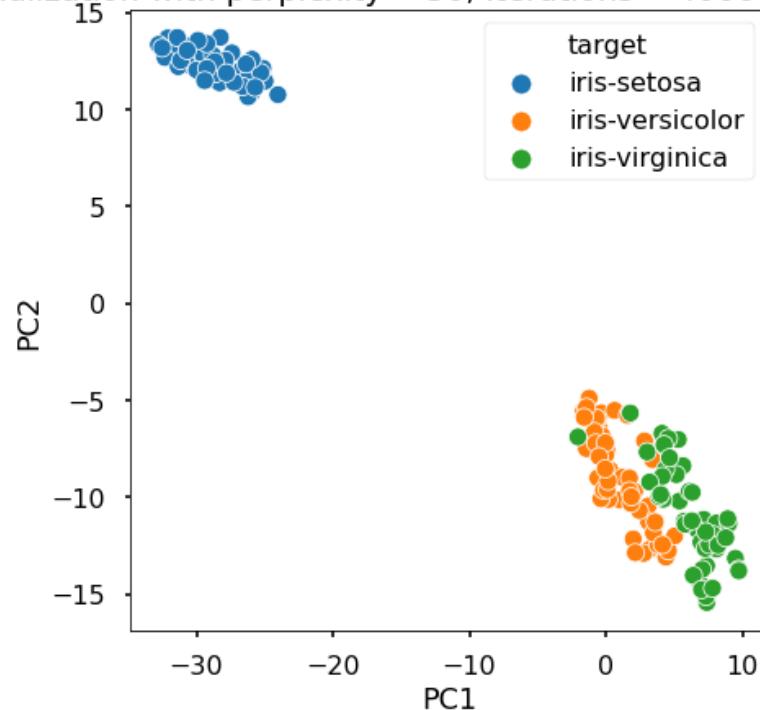
t-SNE visualization with perplexity -- 30, iterations -- 2500 and epsilon -- 200



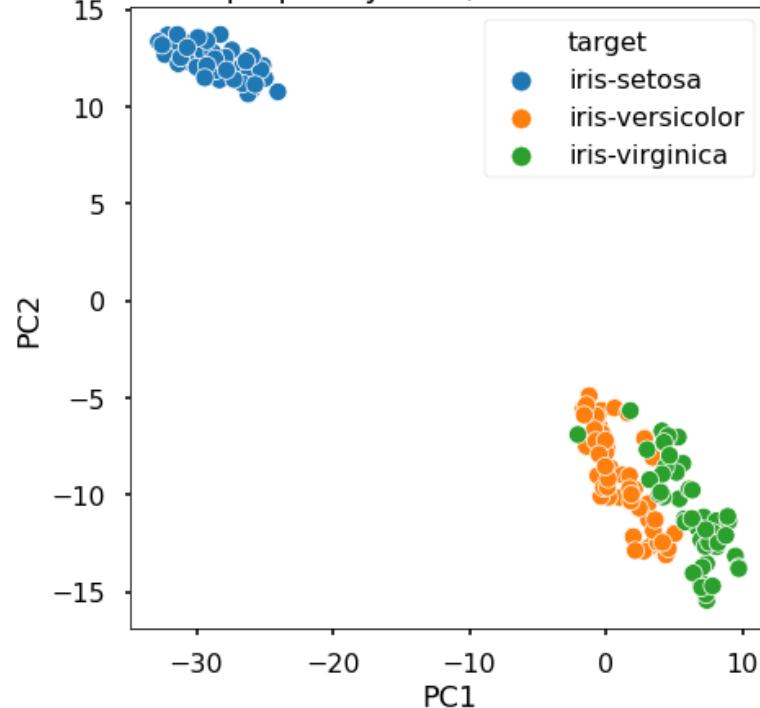
t-SNE visualization with perplexity -- 30, iterations -- 3500 and epsilon -- 200



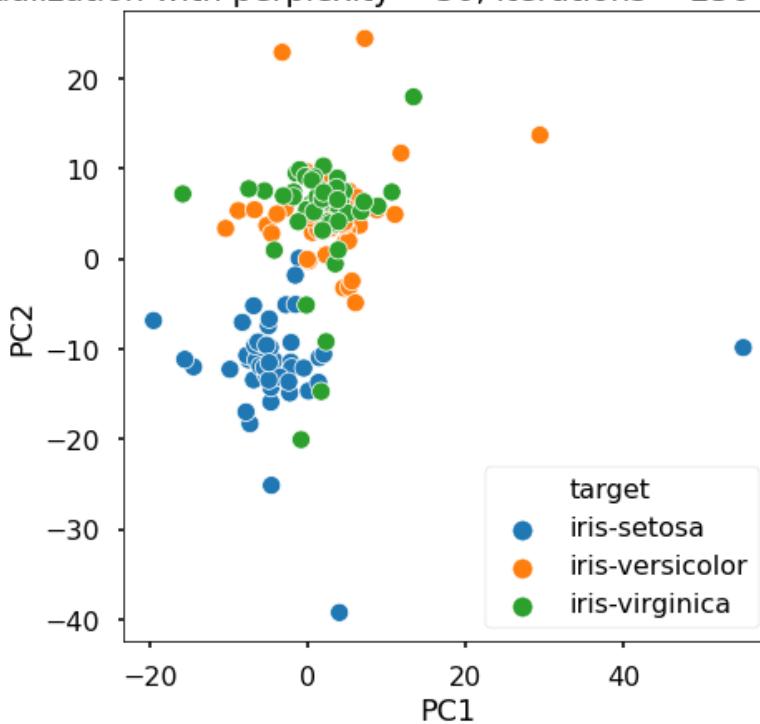
t-SNE visualization with perplexity -- 30, iterations -- 4000 and epsilon -- 200



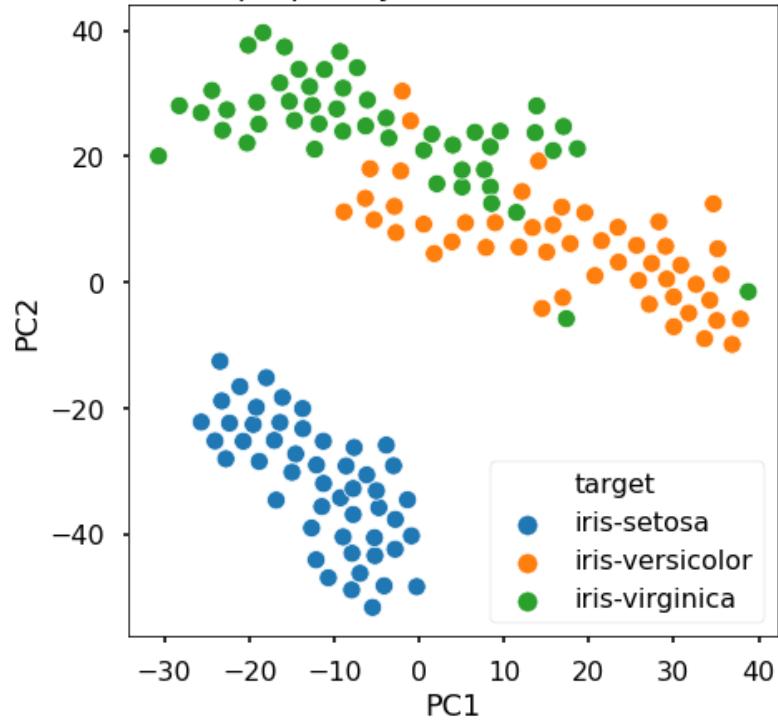
t-SNE visualization with perplexity -- 30, iterations -- 5000 and epsilon -- 200



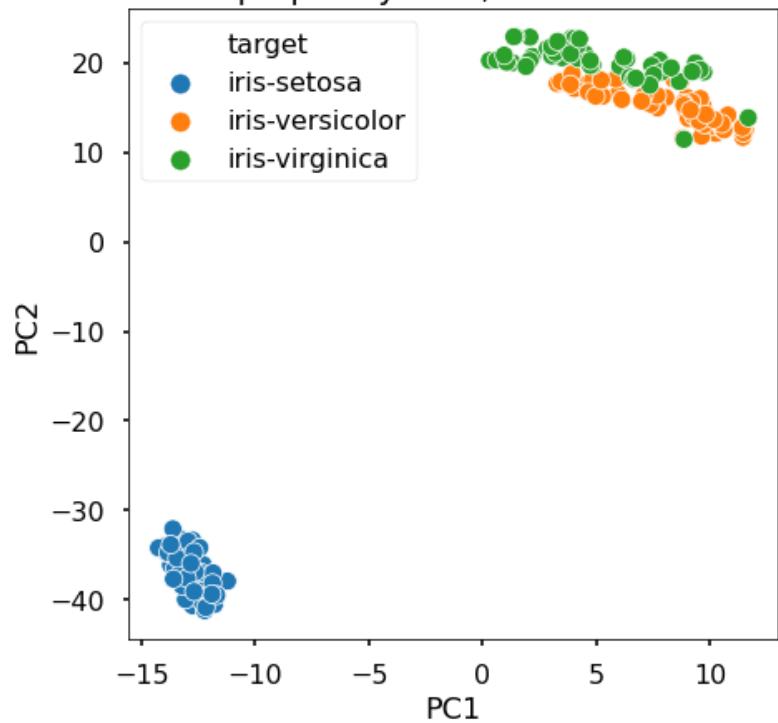
t-SNE visualization with perplexity -- 30, iterations -- 250 and epsilon -- 400



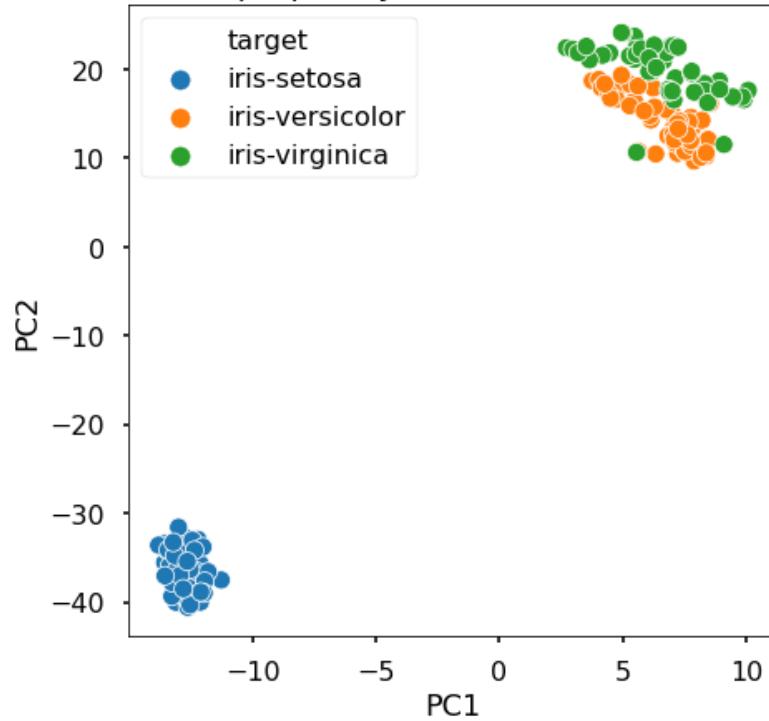
t-SNE visualization with perplexity -- 30, iterations -- 350 and epsilon -- 400



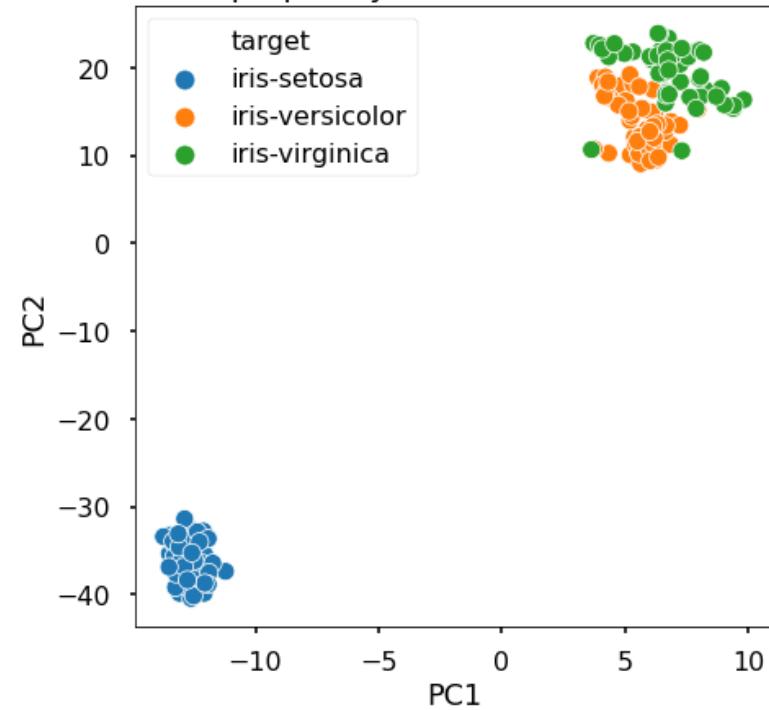
t-SNE visualization with perplexity -- 30, iterations -- 500 and epsilon -- 400



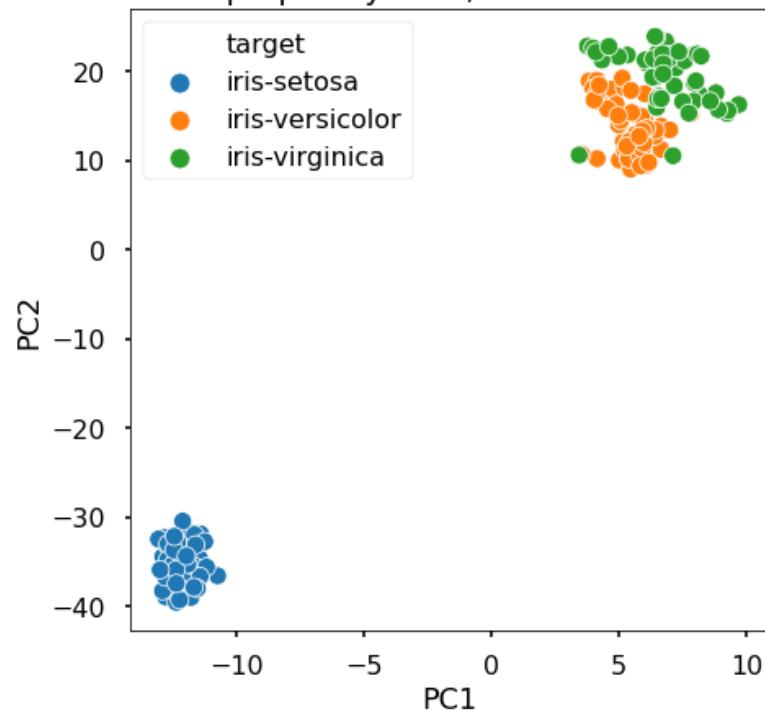
t-SNE visualization with perplexity -- 30, iterations -- 750 and epsilon -- 400



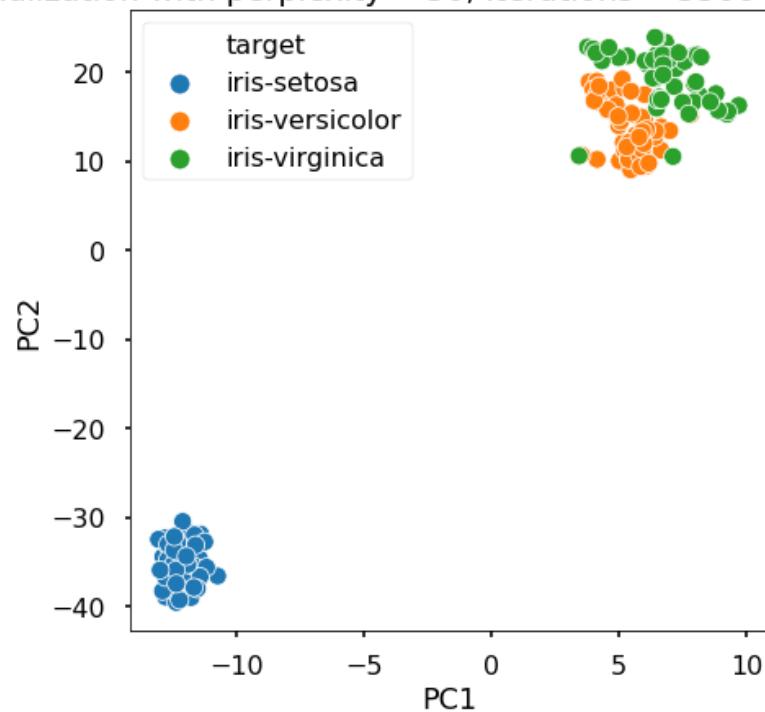
t-SNE visualization with perplexity -- 30, iterations -- 1000 and epsilon -- 400



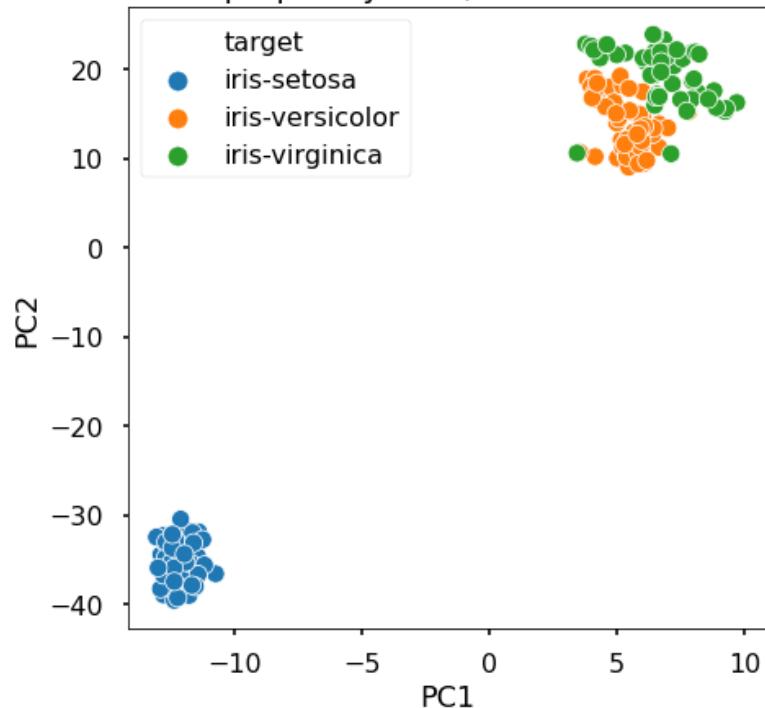
t-SNE visualization with perplexity -- 30, iterations -- 2500 and epsilon -- 400



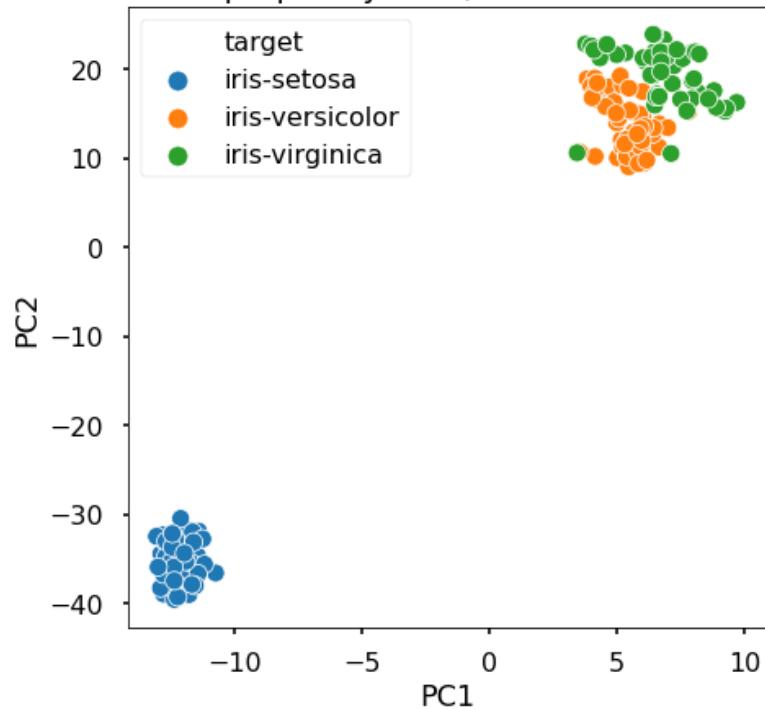
t-SNE visualization with perplexity -- 30, iterations -- 3500 and epsilon -- 400



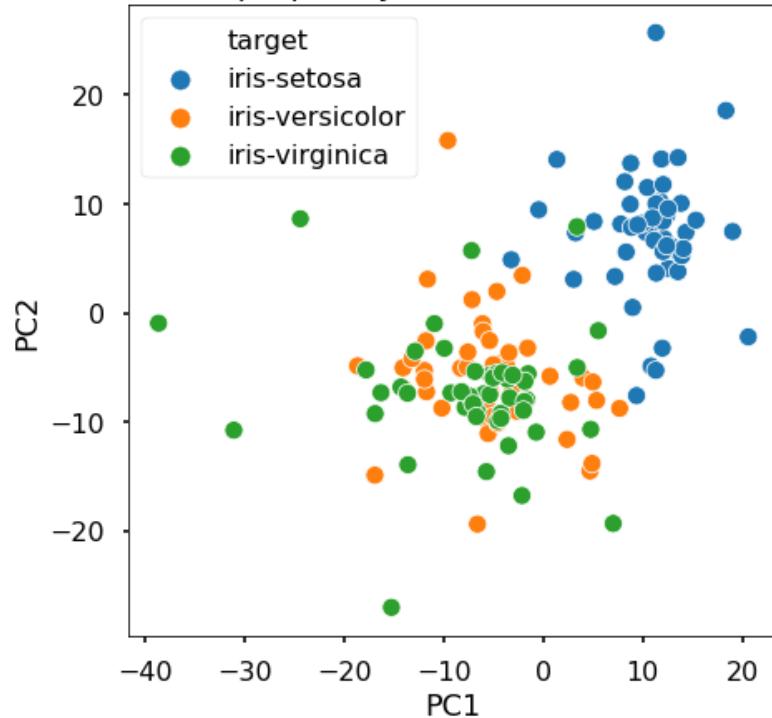
t-SNE visualization with perplexity -- 30, iterations -- 4000 and epsilon -- 400



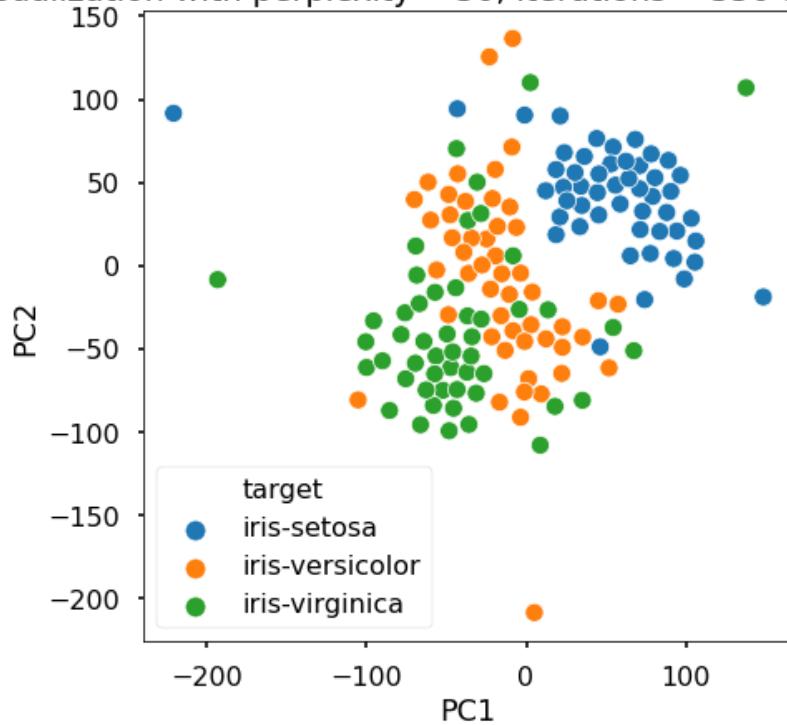
t-SNE visualization with perplexity -- 30, iterations -- 5000 and epsilon -- 400



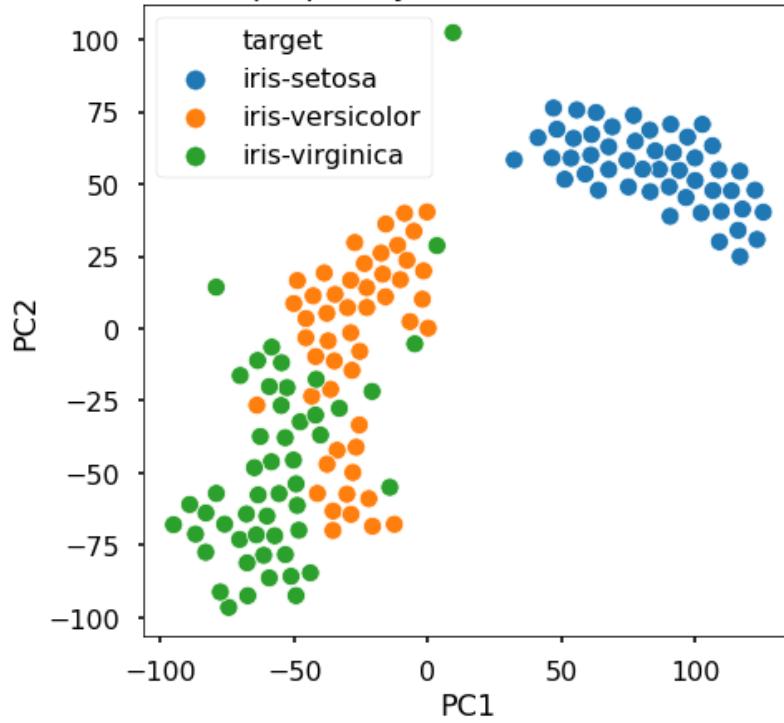
t-SNE visualization with perplexity -- 30, iterations -- 250 and epsilon -- 600



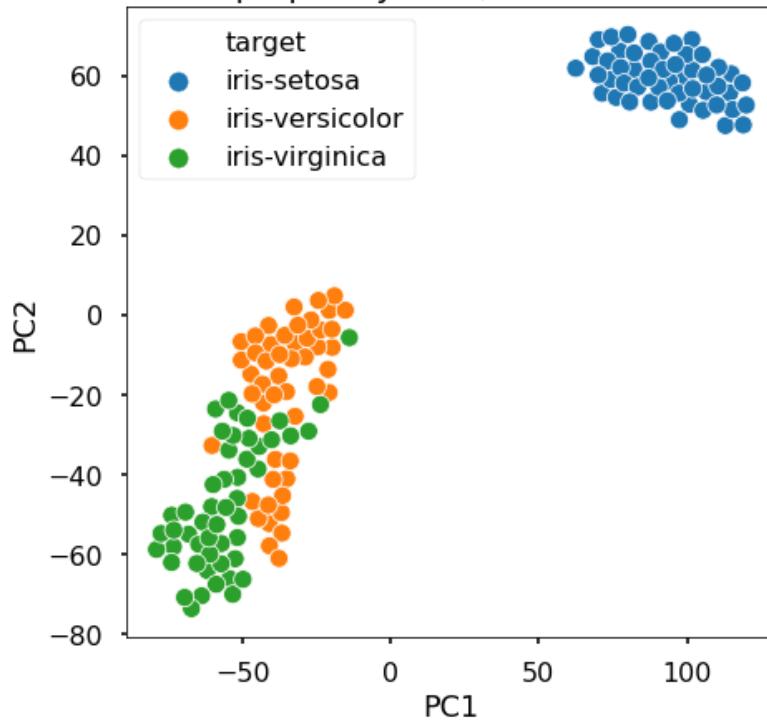
t-SNE visualization with perplexity -- 30, iterations -- 350 and epsilon -- 600



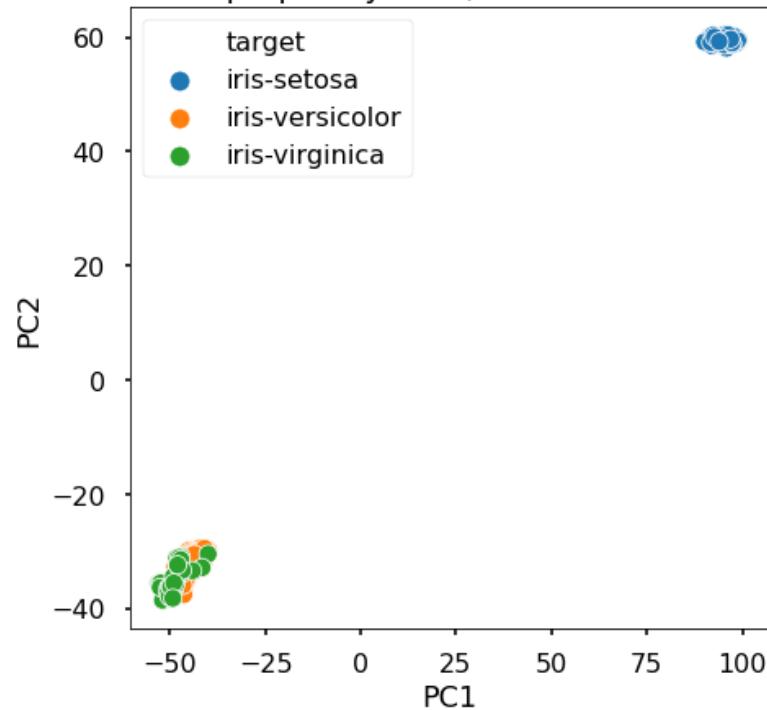
t-SNE visualization with perplexity -- 30, iterations -- 500 and epsilon -- 600



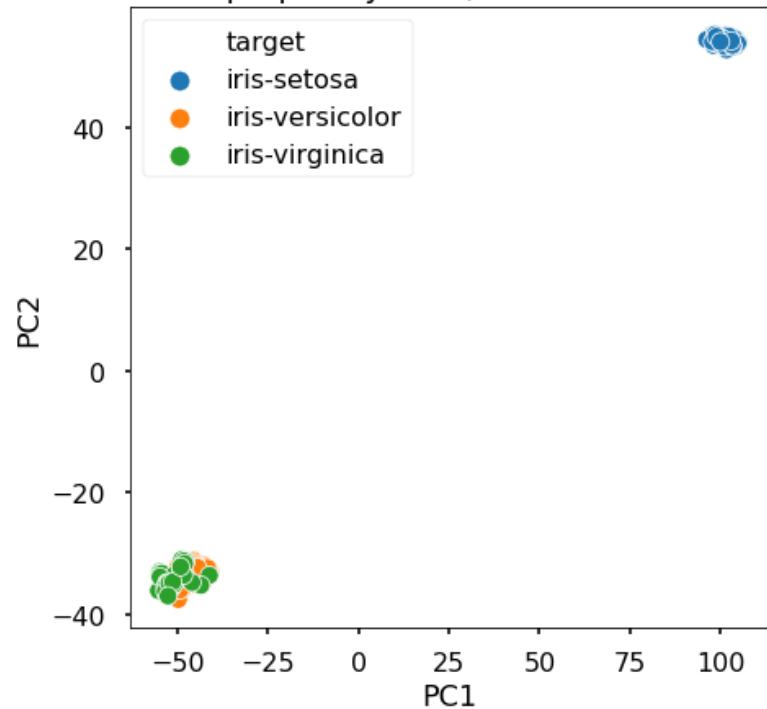
t-SNE visualization with perplexity -- 30, iterations -- 750 and epsilon -- 600



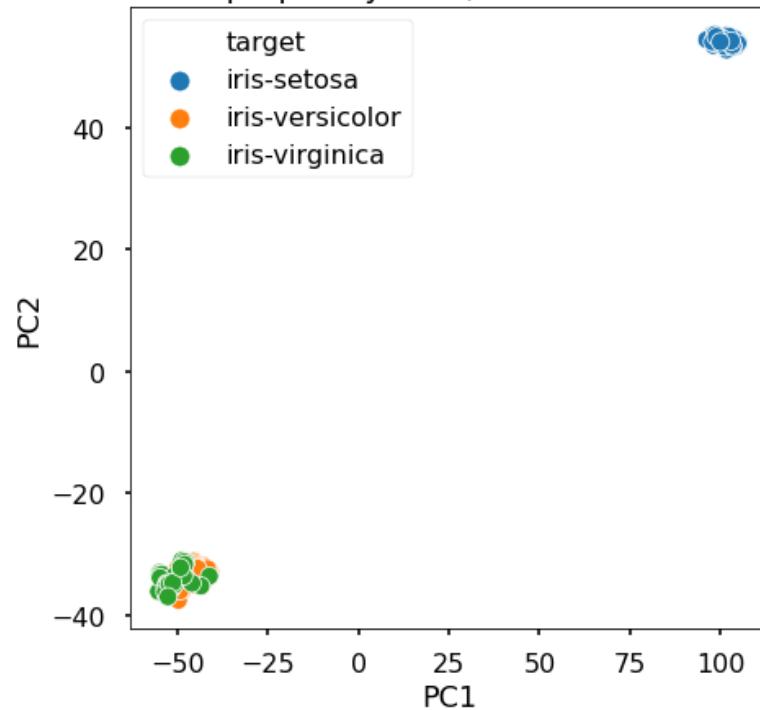
t-SNE visualization with perplexity -- 30, iterations -- 1000 and epsilon -- 600



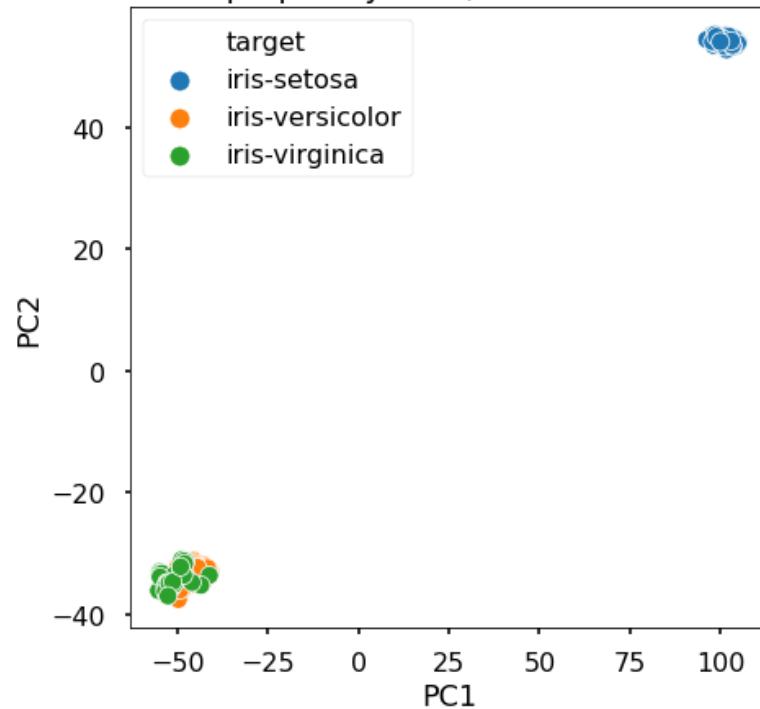
t-SNE visualization with perplexity -- 30, iterations -- 2500 and epsilon -- 600



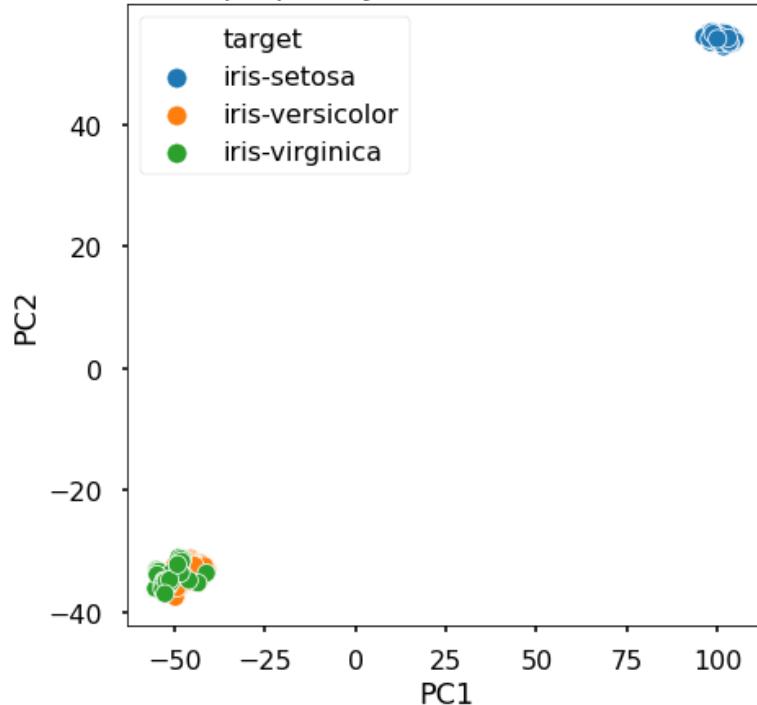
t-SNE visualization with perplexity -- 30, iterations -- 3500 and epsilon -- 600



t-SNE visualization with perplexity -- 30, iterations -- 4000 and epsilon -- 600



t-SNE visualization with perplexity -- 30, iterations -- 5000 and epsilon -- 600



IRIS:CASE-III-Multiple_runs

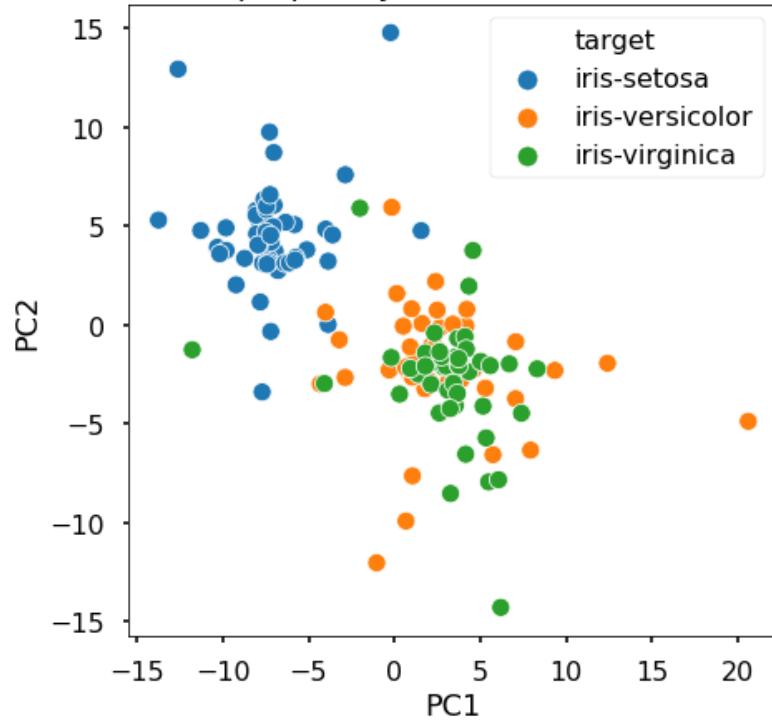
Running t-SNE on a range of iterations of fixed perplexity and Learning rate(epsilon) with embedding initialization as 'random'

```
In [80]: iterations = [250,350,500,750,1000,2500,3500,4000,5000]

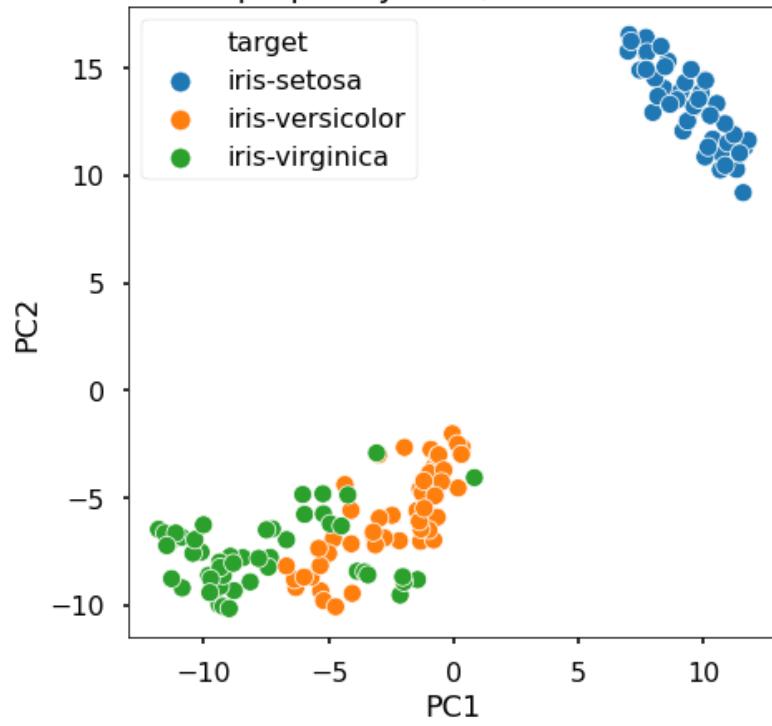
for idx in range(len(iterations)):
    tsne_iris_1 = TSNE(n_components=2,perplexity=30,learning_rate=200,n_iter=iterations[idx],init='random',n_jobs=-1)
    iris1_tsne_pcmps = pd.DataFrame(tsne_iris_1.fit_transform(iris_stand_df.iloc[:,0:4]),columns=['PC1','PC2'])
    iris1_tsne_pcmps = pd.concat([iris1_tsne_pcmps,iris_stand_df['target']],axis=1)
    with plt.style.context('seaborn-poster'):
        plt.figure(figsize=(7,7))
        sns.scatterplot(data=iris1_tsne_pcmps,x='PC1',y='PC2',hue='target')
        plt.title("t-SNE visualization with perplexity -- {0}, iterations -- {1} and epsilon -- {2} ".format(30,iterations[idx],600))

    plt.show()
```

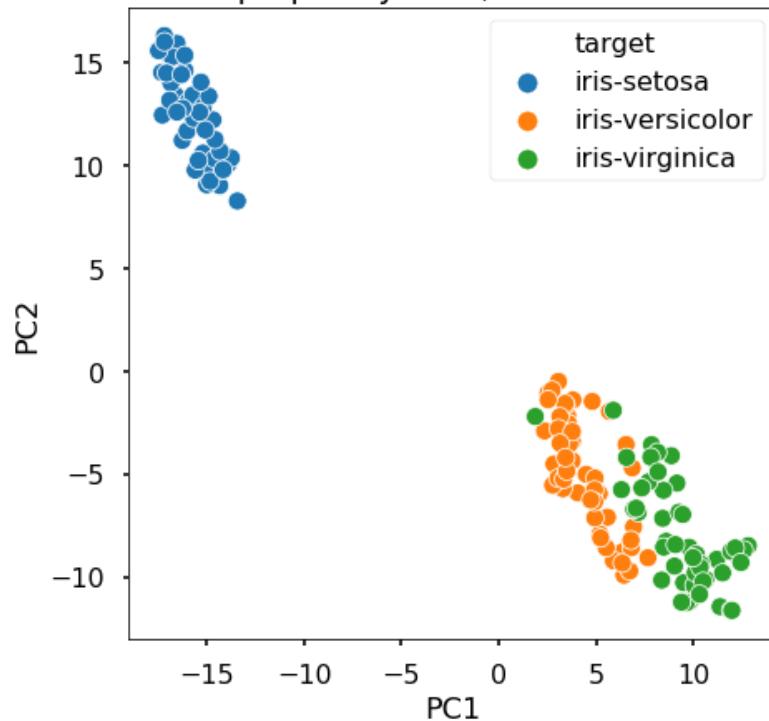
t-SNE visualization with perplexity -- 30, iterations -- 250 and epsilon -- 200



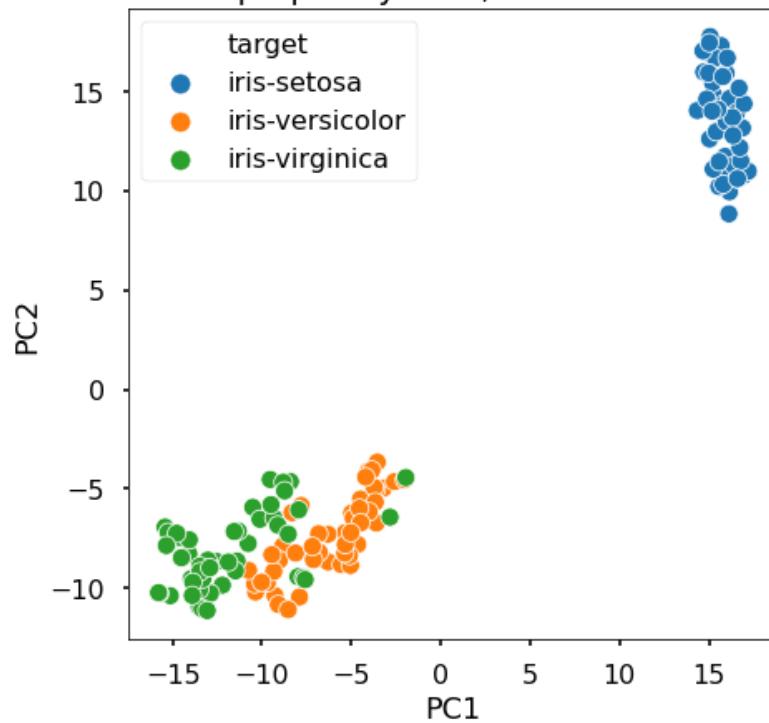
t-SNE visualization with perplexity -- 30, iterations -- 350 and epsilon -- 200



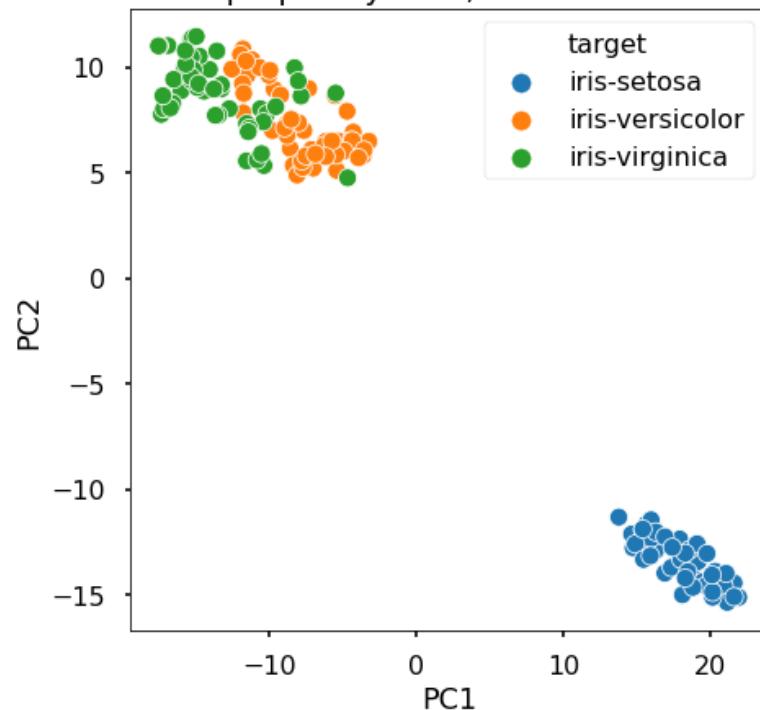
t-SNE visualization with perplexity -- 30, iterations -- 500 and epsilon -- 200



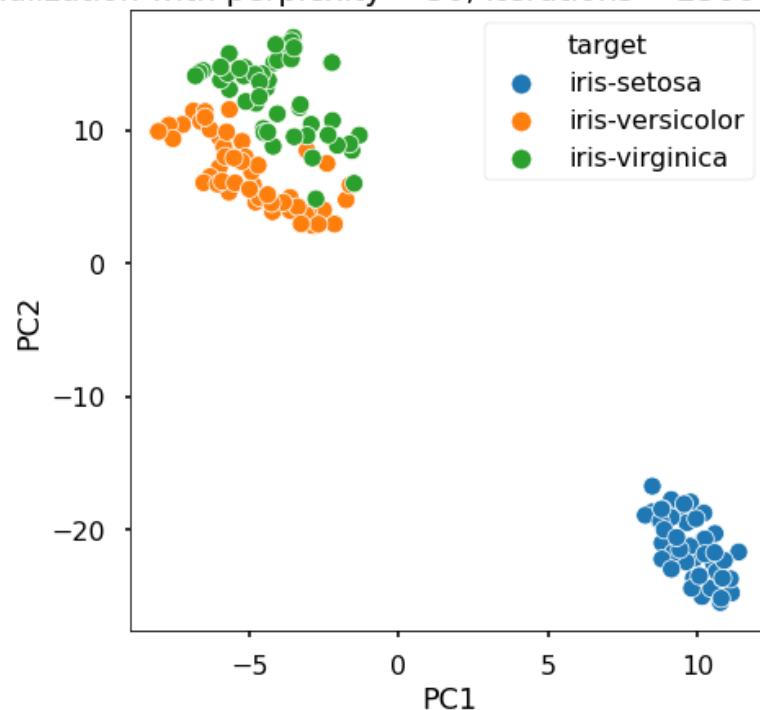
t-SNE visualization with perplexity -- 30, iterations -- 750 and epsilon -- 200



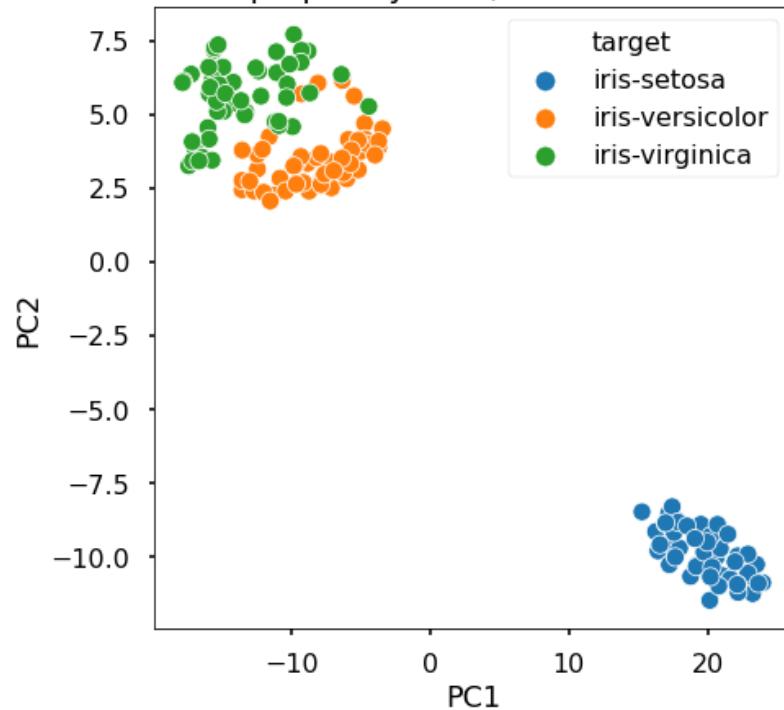
t-SNE visualization with perplexity -- 30, iterations -- 1000 and epsilon -- 200



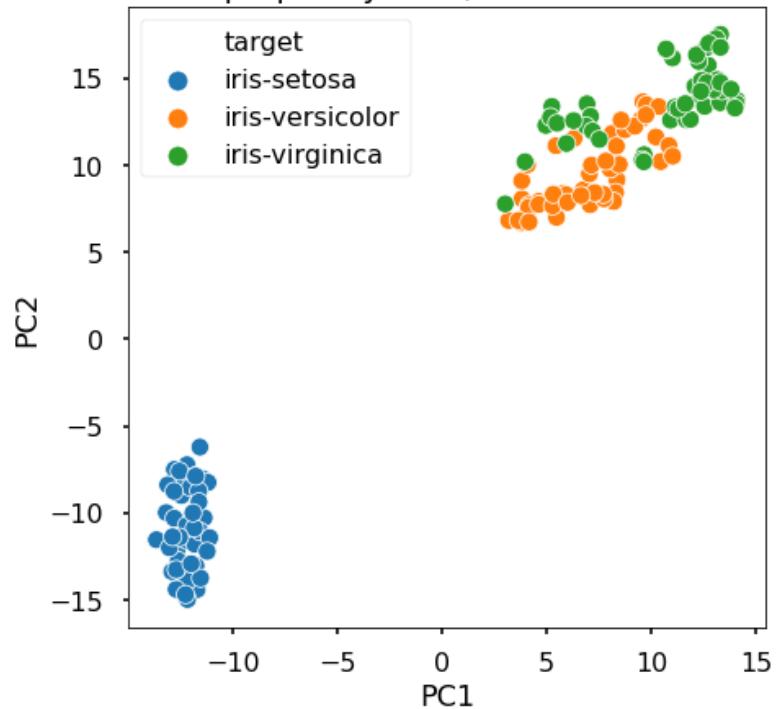
t-SNE visualization with perplexity -- 30, iterations -- 2500 and epsilon -- 200



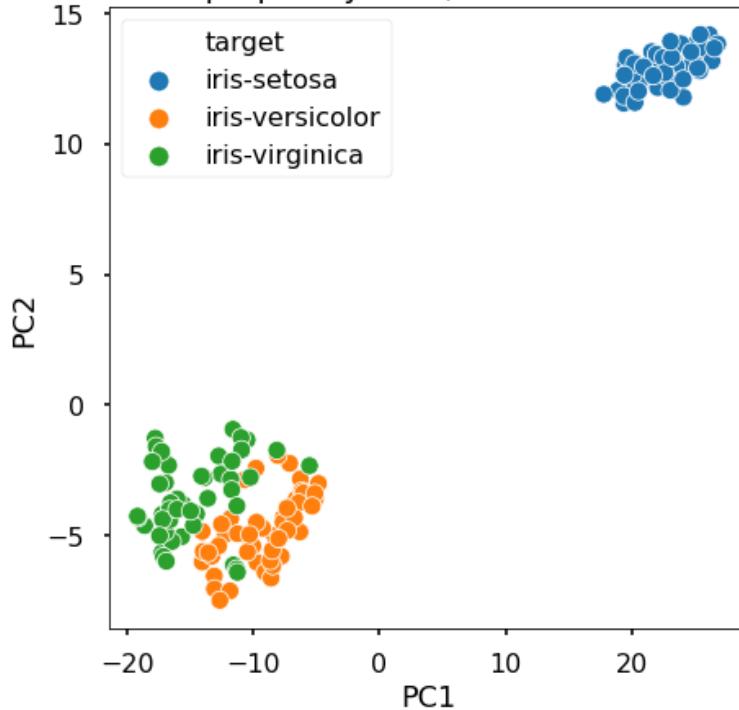
t-SNE visualization with perplexity -- 30, iterations -- 3500 and epsilon -- 200



t-SNE visualization with perplexity -- 30, iterations -- 4000 and epsilon -- 200



t-SNE visualization with perplexity -- 30, iterations -- 5000 and epsilon -- 200

*IRIS:Result_of_multiple_runs*

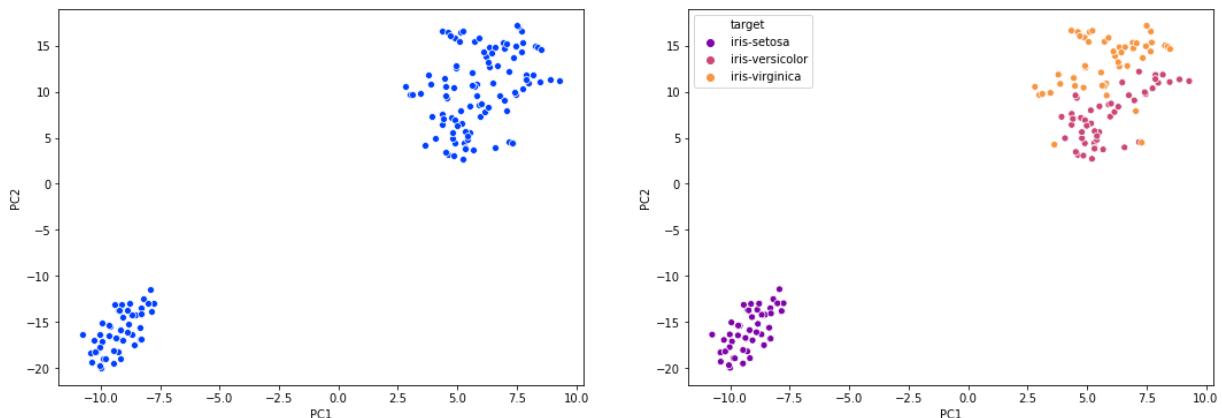
```
In [86]: tsne_iris_1 = TSNE(n_components=2,perplexity=30,learning_rate=200,n_iter=5000,init='
```

```
In [87]: iris_tsne_cmps = pd.concat([pd.DataFrame(tsne_iris_1.fit_transform(iris_stand_df.iloc[:, :-1]), columns=['PC1', 'PC2']), pd.DataFrame(iris_stand_df.target, columns=['target'])], axis=1)
iris_tsne_cmps.head()
```

Out[87]:

	PC1	PC2	target
0	-9.326900	-16.767998	iris-setosa
1	-9.125493	-13.043600	iris-setosa
2	-8.535302	-14.208212	iris-setosa
3	-8.310025	-13.506122	iris-setosa
4	-9.926417	-17.123808	iris-setosa

```
In [89]: with plt.style.context('seaborn-bright'):
    fig,ax = plt.subplots(nrows=1,ncols=2,figsize=(18,6))
    sns.scatterplot(data=iris_tsne_cmps,x='PC1',y='PC2',ax=ax[0])
    sns.scatterplot(data=iris_tsne_cmps,x='PC1',y='PC2',hue='target',ax=ax[1],palette='Set1')
```



Other_method

IRIS : t-SNE with 'pca' initialization embedding, Nearest neighbors error method as 'exact' and distance metric as 'minkowski'

```
In [92]: import scipy

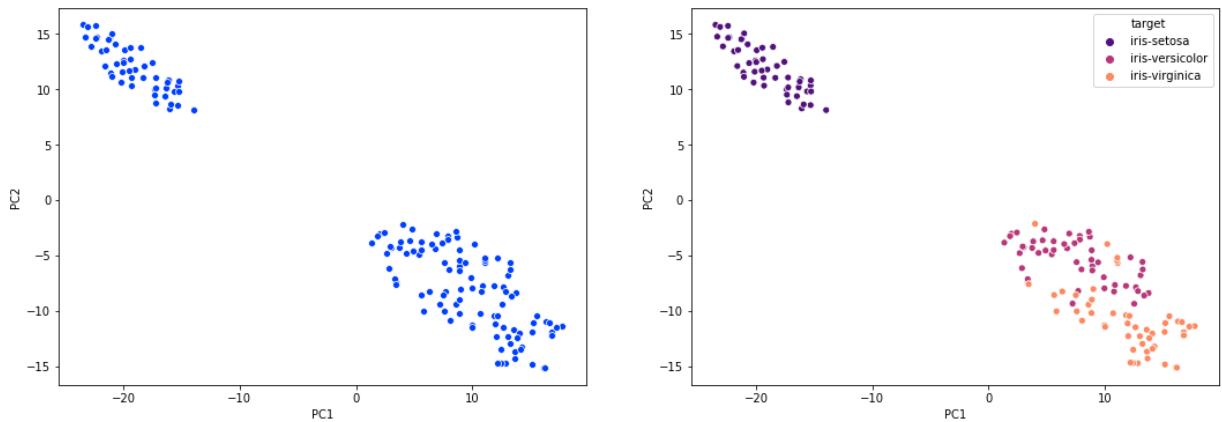
tsne_iris_rand = TSNE(n_components=2,
                      perplexity=30,
                      learning_rate=200,
                      n_iter=5000,
                      init='pca',           ## Good with the global capturing
                      method='exact',       ## This method is slower and not s
                      metric=scipy.spatial.minkowski_distance,
                      n_jobs=-1)
```



```
In [93]: iris_tsne_cmps = pd.concat([pd.DataFrame(tsne_iris_rand.fit_transform(iris_stand_df),
                                                pd.DataFrame(iris_stand_df.target,columns=['target'])],axis=1)
iris_tsne_cmps.head()
```

	PC1	PC2	target
0	-20.068819	12.580353	iris-setosa
1	-15.324358	10.337151	iris-setosa
2	-17.385283	9.968454	iris-setosa
3	-16.503378	9.369566	iris-setosa
4	-19.959255	13.517425	iris-setosa

```
In [94]: with plt.style.context('seaborn-bright'):
    fig,ax = plt.subplots(nrows=1,ncols=2,figsize=(18,6))
    sns.scatterplot(data=iris_tsne_cmps,x='PC1',y='PC2',ax=ax[0])
    sns.scatterplot(data=iris_tsne_cmps,x='PC1',y='PC2',hue='target',ax=ax[1],palette='Set1')
```



Conclusion

In my analysis with two different dataset I found that t-SNE gives better results when embedding method is 'PCA' inplace of 'random' and nearest

neighbors error method as 'exact' instead of 'barnes_hut'.

The 'barnes_hut' method is much better in terms of execution performance as compared to 'exact' and it is not a good approach with a large dataset. If you want the nearest neighbors error to be less than 3% and your dataset is not very large then better go with 'exact'.