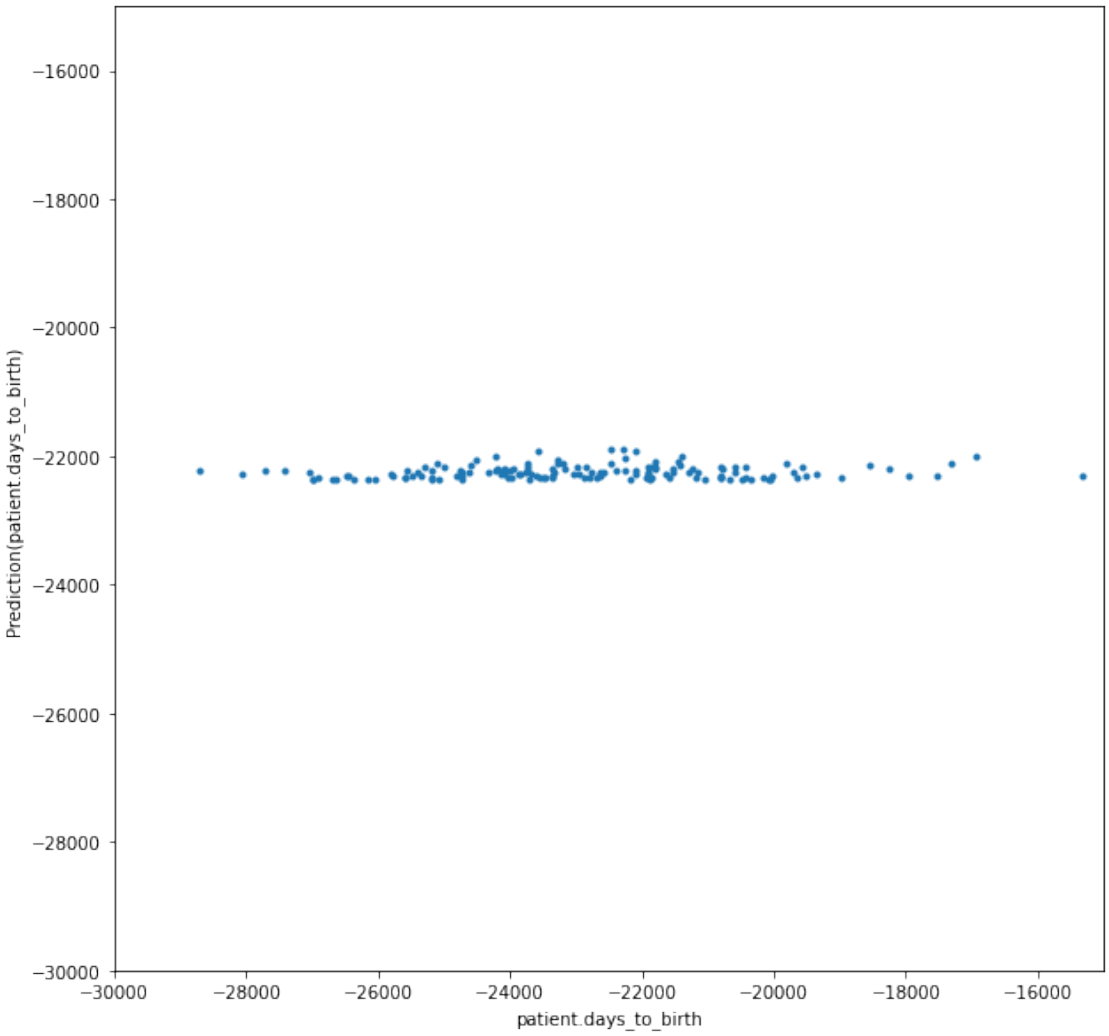
Regresja liniowa i klasyfikacja SVM

March 17, 2025

[1]:
```
# Regresja Liniowa

# Ukryj ostrzeżenia (niekompatybilne wersje bibliotek)
import warnings
warnings.filterwarnings('ignore')

import numpy as np
import pandas as pd

from sklearn.linear_model import LinearRegression
import matplotlib.pyplot as plt

data = pd.read_csv('/home/lpo/CancerProstateSurvival.csv', sep=',')
filtered_data = data[data['patient.days_to_birth'].notnull()]
times = filtered_data['times'].map(int).values
days = filtered_data['patient.days_to_birth'].map(int).values
regression_part = int(len(times) * 0.7)

regression_parameter = times[:regression_part].reshape((-1, 1))
regression_target = days[:regression_part]

model = LinearRegression().fit(regression_parameter, regression_target)

prediction_parameter = times[regression_part:].reshape((-1, 1))
prediction_target = days[regression_part:]

r_2 = model.score(prediction_parameter, prediction_target)
print(f"R^2 = {r_2}")

plt.figure(figsize=(10, 10))
plt.plot(prediction_target, model.predict(prediction_parameter), '.')
ax = plt.gca()
ax.set_ylim([-30000, -15000])
ax.set_xlim([-30000, -15000])
plt.xlabel('patient.days_to_birth')
plt.ylabel('Prediction(patient.days_to_birth)')
plt.show()
```

```
R^2 = -0.0719439601598717
```



[2]:
```
# SVM

from sklearn.svm import SVC

ethnicity = filtered_data['patient.ethnicity'].map(lambda x: 0 if x != x else␣
 ↪1).values
parameters = list(zip(times, days))
regression_part = int(len(times) * 0.3)

regression_parameters = parameters[regression_part:]
regression_target = ethnicity[regression_part:]

model = SVC(kernel='rbf').fit(regression_parameters, regression_target)
```

```
prediction_parameter = parameters[:regression_part]
prediction_target = ethnicity[:regression_part]

prediction = model.predict(prediction_parameter)
total = len(prediction_target)
succeed = len([1 for a, b in zip(prediction, prediction_target) if a == b])
print(f"Prawidłowych predykcji: {succeed / total * 100.0} %")

plt.figure(figsize=(10, 10))
plt.pie([total - succeed, succeed], labels=['Błędna klasyfikacja', 'Poprawna␣
 ↪klasyfikacja'])
plt.show()
```

```
Prawidłowych predykcji: 37.67123287671233 %
```