

Incidenti Stradali

Alessandro Beranti

Università degli Studi di Milano Statale

25 Febbraio 2020

Introduzione

Obiettivo

Stabilire che tipo di mezzo ha investito la vittima

Caratteristiche iniziali

A partire dal file fornito ho eseguito diversi studi considerando di volta in volta diversi fattori:

- ▶ Caratteristiche delle ossa rotte
- ▶ Algoritmi di apprendimento
- ▶ Pre-processing dei dati tramite:
 - ▶ StandardScaler
 - ▶ MinMaxScaler
 - ▶ RobustScaler

Risultati ottenuti

Da ogni studio sono conseguite diverse tabelle che tengono conto dei fattori precedentemente descritti.

Modelli Apprendimento

Modelli usati

Per compiere lo studio sono stati usati diversi modelli di apprendimento:

- ▶ Support Vector Machine (SVC)
- ▶ Decision Tree (DT)
- ▶ Random Forest (RF)
- ▶ GaussianNB (NB)
- ▶ LinearDiscriminantAnalysis (LD)
- ▶ MLPClassifier (MLP)

Componenti usate

Totali:

- ▶ Sesso, Anni, Peso, Altezza, Tot Testa, Tot Torace, Tot Addome, Tot Scheletro

Totali_BMI:

- ▶ Sesso, Anni, Peso, Altezza, Bmi, Tot Testa, Tot Torace, Tot Addome, Tot Scheletro

Totali_DATA:

- ▶ Data, Sesso, Anni, Peso, Altezza, Tot Testa, Tot Torace, Tot Addome, Tot Scheletro

Componenti usate

Totali_BMI_DATA:

- ▶ Data, Sesso, Anni, Peso, Altezza, Bmi, Tot Testa, Tot Torace, Tot Addome, Tot Scheletro

Details, composto da:

▶ Testa

- ▶ Neurocranio
- ▶ Splancnocranio
- ▶ Telencefalo
- ▶ Cervelletto
- ▶ Tronco-encefalico

▶ Torace

- ▶ Polmoni
- ▶ Trachea/bronchi
- ▶ Cuore
- ▶ Aorta-toracica
- ▶ Diaframma

Componenti usate

▶ Addome

- ▶ Fegato
- ▶ Milza
- ▶ Aorta-addominale
- ▶ Reni
- ▶ Mesentere

▶ Scheletro

- ▶ Rachide-cervicale
- ▶ Rachide-toracico
- ▶ Rachide-lombare
- ▶ bacino-e-sacro
- ▶ Complesso-
sterno/claveo/costale

Analizziamo i risultati

- ▶ PCA e TSNE sono gli algoritmi usati per ridurre la dimensionalità
- ▶ Il numero a fianco indica a quante dimensioni si è scesi
- ▶ Details ha una dimensione iniziale di 20
- ▶ Il punteggio sta ad indicare la percentuale di casi in cui l'algoritmo ha predetto correttamente di che mezzo si tratta

Analizziamo i risultati

StandardScaler

	Totali	Totali_BMI	Totali_DATA	Totali_DATA_and_BMI
SVC	0.66	0.66	0.77	0.73
DT	0.56	0.56	0.61	0.63
RF	0.64	0.63	0.57	0.59
NB	0.69	0.68	0.72	0.75
LD	0.71	0.68	0.72	0.72
MLP	0.69	0.69	0.74	0.72

	Details	Details PCA_5	Details PCA_10	Details PCA_13	Details PCA_15	Details TSNE_13	Details TSNE_15
SVC	0.68	0.63	0.68	0.67	0.71	0.51	0.53
DT	0.68	0.60	0.67	0.64	0.57	0.43	0.50
RF	0.67	0.59	0.69	0.70	0.63	0.51	0.55
NB	0.63	0.61	0.66	0.67	0.65	0.56	0.57
LD	0.63	0.64	0.60	0.59	0.59	0.48	0.49
MLP	0.63	0.62	0.61	0.62	0.65	0.39	0.47

Analizziamo i risultati

MinMaxScaler

	Totali	Totali_BMI	Totali_DATA	Totali_DATA_and_BMI
SVC	0.66	0.62	0.73	0.74
DT	0.61	0.55	0.65	0.63
RF	0.60	0.62	0.62	0.61
NB	0.69	0.68	0.72	0.75
LD	0.71	0.68	0.72	0.72
MLP	0.62	0.64	0.71	0.76

	Details	Details PCA_5	Details PCA_10	Details PCA_13	Details PCA_15	Details TSNE_13	Details TSNE_15
SVC	0.74	0.70	0.64	0.67	0.70	0.58	0.42
DT	0.70	0.61	0.60	0.56	0.59	0.51	0.56
RF	0.67	0.65	0.66	0.71	0.67	0.46	0.52
NB	0.63	0.67	0.65	0.67	0.67	0.57	0.58
LD	0.63	0.65	0.61	0.60	0.62	0.43	0.39
MLP	0.65	0.64	0.64	0.60	0.63	0.51	0.43

Analizziamo i risultati

RobustScaler

	Totali	Totali_BMI	Totali_DATA	Totali_DATA_and_BMI
SVC	0.66	0.68	0.77	0.72
DT	0.59	0.60	0.62	0.65
RF	0.62	0.63	0.56	0.61
NB	0.69	0.68	0.72	0.75
LD	0.71	0.68	0.72	0.72
MLP	0.70	0.67	0.70	0.66

	Details	Details_PCA_5	Details_PCA_10	Details_PCA_13	Details_PCA_15	Details_TSNE_13	Details_TSNE_15
SVC	0.68	0.63	0.63	0.66	0.66	0.56	0.47
DT	0.65	0.63	0.61	0.64	0.63	0.45	0.56
RF	0.65	0.65	0.64	0.64	0.66	0.52	0.49
NB	0.63	0.67	0.67	0.67	0.65	0.54	0.56
LD	0.63	0.63	0.60	0.64	0.67	0.43	0.42
MLP	0.56	0.63	0.60	0.65	0.60	0.49	0.46

Analizziamo i risultati

Senza Scaler

	Details PCA_5	Details PCA_10	Details PCA_13	Details PCA_15	Details TSNE_13	Details TSNE_15
SVC	0.65	0.66	0.68	0.66	0.57	0.54
DT	0.60	0.55	0.59	0.51	0.50	0.52
RF	0.66	0.64	0.60	0.64	0.53	0.55
NB	0.67	0.64	0.66	0.65	0.52	0.55
LD	0.64	0.62	0.58	0.62	0.47	0.51
MLP	0.66	0.65	0.60	0.58	0.43	0.53

In generale non c'è un aumento di prestazioni tra i vari scaler

Conclusioni

	StandardScaler	MinMaxScaler	RobustScaler	No_Scaler
Totali	0.713	0.713	0.713	NaN
Totali_BMI	0.692	0.689	0.689	NaN
Totali_DATA	0.774	0.735	0.774	NaN
Totali_DATA_BMI	0.751	0.767	0.751	NaN
Details	0.685	0.740	0.685	NaN
Details_PCA_5	0.647	0.701	0.670	0.670
Details_PCA_10	0.699	0.662	0.670	0.662
Details_PCA_13	0.701	0.710	0.670	0.685
Details_PCA_15	0.715	0.700	0.678	0.661
Details_TSNE_13	0.561	0.584	0.568	0.572
Details_TSNE_15	0.578	0.586	0.562	0.554

Le migliori prestazioni si ottengono usando Totali con la data senza praticare una riduzione della dimensionalità