

Постановка задачи:

Обучиться структуре байесовской сети, которая должна уметь прогнозировать G1, G2, G3, оценить параметры каждой полученной структуры, оценить ошибку прогноза каждой из получившихся сетей; каждую получившуюся сеть изобразить в виде направленного графа с подписанными вершинами. Сравнить все полученные результаты с какой-либо моделью машинного обучения.

1. Данные

Данные представляют собой описание студентов математического курса (student-mat.csv) и курса португальского языка (student-por.csv).

2. Предварительная обработка

Исходные данные student-mat.csv:

	school	sex	age	address	famsize	Pstatus	Medu	Fedu	Mjob	Fjob	...	famrel	freetime	goout	Dalc	Walc	health	absences	G1	G2	G3
0	GP	F	18	U	GT3	A	4	4	at_home	teacher	...	4	3	4	1	1	3	6	5	6	6
1	GP	F	17	U	GT3	T	1	1	at_home	other	...	5	3	3	1	1	3	4	5	5	6
2	GP	F	15	U	LE3	T	1	1	at_home	other	...	4	3	2	2	3	3	10	7	8	10
3	GP	F	15	U	GT3	T	4	2	health	services	...	3	2	2	1	1	5	2	15	14	15
4	GP	F	16	U	GT3	T	3	3	other	other	...	4	3	2	1	2	5	4	6	10	10

5 rows × 33 columns

Данные проверены на наличие пропущенных значений, которые не были обнаружены.

Также для уменьшения количества категорий в некоторых категориальных признаках проведено объединение схожих значений. Числовые признаки сделаны категориальными путем дискретизации и заданы метки. Преобразования выполнены для признаков "failures", "studytime", "absences", "G1", "G2", "G3".

Результат:

	school	sex	age	address	famsize	Pstatus	Medu	Fedu	Mjob	Fjob	...	famrel	freetime	goout	Dalc	Walc	health	absences	G1	G2	G3
0	GP	F	18	U	GT3	A	4	4	at_home	teacher	...	4	3	4	1	1	3	Low-absence	Fail	Fail	Fail
1	GP	F	17	U	GT3	T	1	1	at_home	other	...	5	3	3	1	1	3	Low-absence	Fail	Fail	Fail
2	GP	F	15	U	LE3	T	1	1	at_home	other	...	4	3	2	2	3	3	High-absence	Fail	Fail	Pass
3	GP	F	15	U	GT3	T	4	2	health	services	...	3	2	2	1	1	5	Low-absence	Pass	Pass	Pass
4	GP	F	16	U	GT3	T	3	3	other	other	...	4	3	2	1	2	5	Low-absence	Fail	Pass	Pass
...
390	MS	M	20	U	LE3	A	2	2	services	services	...	5	5	4	4	5	4	High-absence	Fail	Fail	Fail
391	MS	M	17	U	LE3	T	3	1	services	services	...	2	4	5	3	4	2	Low-absence	Pass	Pass	Pass
392	MS	M	21	R	GT3	T	1	1	other	other	...	5	5	3	3	3	3	Low-absence	Pass	Fail	Fail
393	MS	M	18	R	LE3	T	3	2	services	other	...	4	4	1	3	4	5	No-absence	Pass	Pass	Pass
394	MS	M	19	U	LE3	T	1	1	other	at_home	...	3	2	3	3	3	5	Low-absence	Fail	Fail	Fail

Затем все переменные переведены в числовые.

Для прогнозируемых признаков "G1", "G2", "G3" значение 0 соответствует Fail (при условии, что балл ниже 10), значение 1 – Pass (при условии, что балл > или = 10).

В анализ не включались чувствительные переменные, такие как 'age', 'school', 'sex', 'Mjob', 'Fjob', 'reason', 'guardian'.

Данные, которые использованы для анализа:

	address	famsize	Pstatus	Medu	Fedu	traveltime	studytime	failures	schoolsup	famsup	...	famrel	freetime	goout	Dalc	Walc	health	absences	G1	G2	G3
0	1	0	0	4	4	2	1	1	1	0	...	4	3	4	1	1	3	1	0	0	0
1	1	0	1	1	1	1	1	1	0	1	...	5	3	3	1	1	3	1	0	0	0
2	1	1	1	1	1	1	1	0	1	0	...	4	3	2	2	3	3	0	0	0	1
3	1	0	1	4	2	1	0	1	0	1	...	3	2	2	1	1	5	1	1	1	1
4	1	0	1	3	3	1	1	1	0	1	...	4	3	2	1	2	5	1	0	1	1
...
390	1	1	0	2	2	1	1	0	0	1	...	5	5	4	4	5	4	0	0	0	0
391	1	1	1	3	1	2	1	1	0	0	...	2	4	5	3	4	2	1	1	1	1
392	0	0	1	1	1	1	1	0	0	0	...	5	5	3	3	3	3	1	1	0	0
393	0	1	1	3	2	3	1	1	0	0	...	4	4	1	3	4	5	2	1	1	1
394	1	1	1	1	1	1	1	1	0	0	...	3	2	3	3	3	5	1	0	0	0

395 rows x 26 columns

3. Ход работы:

Для структурного обучения байесовской сети использована библиотека `bnlearn`. Настройкой по умолчанию в `bnlearn` для структурного обучения является метод поиска холма (`hillclimbsearch`) и оценка BIC. Применение такого метода не показало связей между вершинами графа. Также изучение структуры с более чем 15 узлами вычислительно невозможно при исчерпывающем поиске (`exhaustivesearch`).

Поэтому в ходе работы рассмотрены Chow-liu и Tree-augmented Naive Bayes (TAN).

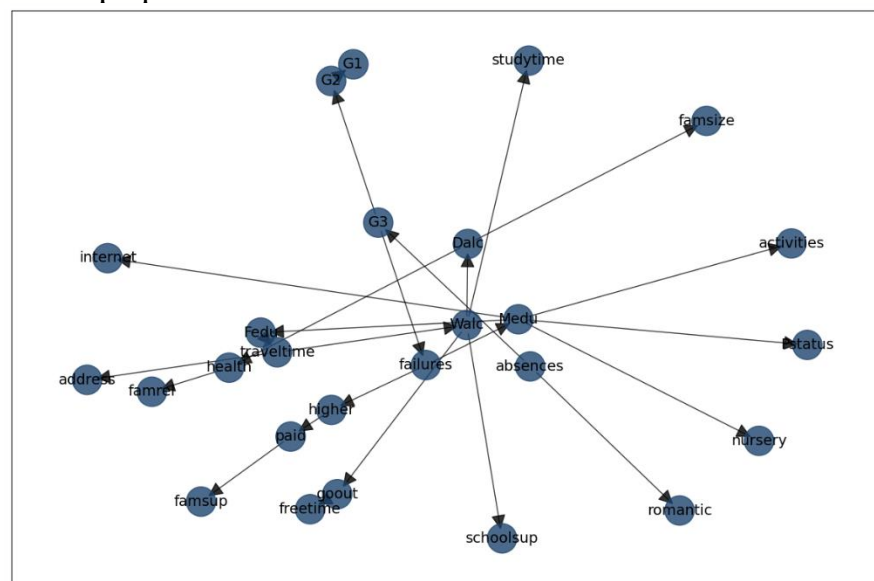
Данные поделены на обучающую выборку (`train`, 0.9) и тестовую (`test`, 0.1). Используя обучающую выборку, производится структурное обучение байесовской сети.

Chow-liu

В качестве `root_node` взят `'absences'`. Вычислены оценки структуры `['k2', 'bds', 'bic', 'bdeu']` для сравнения моделей (чем выше, тем лучше). В результате для `scoretype` взят `'k2'`, т.к.:

```
'structure_scores': {'k2': -11728.583628973223,
  'bds': -12692.001852486801,
  'bic': -17315.99000937574,
  'bdeu': -12449.40033929082}}
```

Полученный граф:



Далее проведен parameter learning и построены CPD. Таблицы условной вероятности (CPD) описывают статистические отношения между каждым узлом и его родителями. В данном задании использована Байесовская оценка.

Пример полученного результата (другие CPD представлены в приложении):

CPD of absences:

+-----+-----+		
absences(0)	0.302091	
+-----+-----+		
absences(1)	0.375154	
+-----+-----+		
absences(2)	0.322755	
+-----+-----+		

CPD of G3:

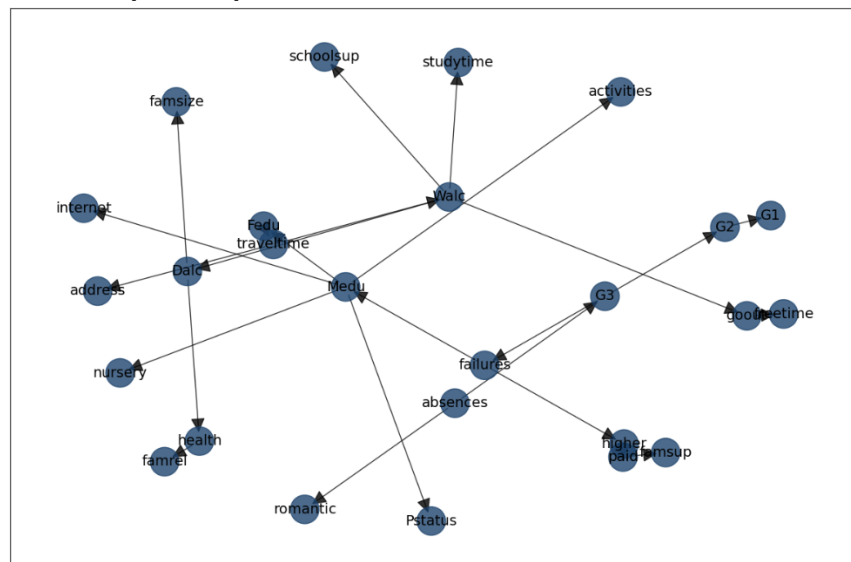
+-----+-----+-----+-----+				
absences	absences(0)	absences(1)	absences(2)	
+-----+-----+-----+-----+				
G3(0)	0.49022801302931596	0.4026229508196722	0.49085365853658536	
+-----+-----+-----+-----+				
G3(1)	0.509771986970684	0.5973770491803279	0.5091463414634146	
+-----+-----+-----+-----+				

Для оценки качества построенной модели прогноза G1 построена матрица ошибок и отчет с указанием accuracy, precision, recall и F-меры:

[[7 2] [2 29]]				
	precision	recall	f1-score	support
0	0.78	0.78	0.78	9
1	0.94	0.94	0.94	31
accuracy			0.90	40
macro avg	0.86	0.86	0.86	40
weighted avg	0.90	0.90	0.90	40

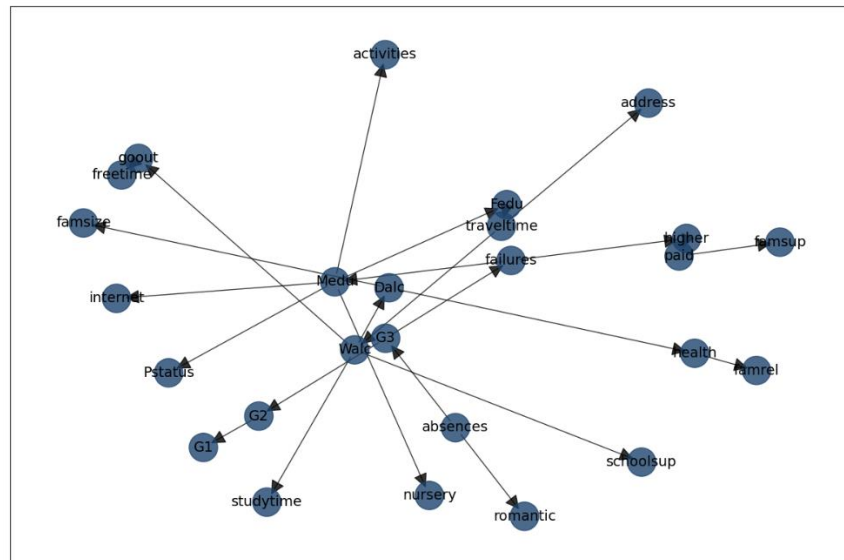
Посмотрев на полученные значения, например, на f1, можно заметить, что модель хорошо предсказывает и тех, кто сдал, и тех, кто не сдал.

Аналогично построен прогноз и оценено качество для G2:



[[7 2] [0 31]]				
	precision	recall	f1-score	support
0	1.00	0.78	0.88	9
1	0.94	1.00	0.97	31
accuracy			0.95	40
macro avg	0.97	0.89	0.92	40
weighted avg	0.95	0.95	0.95	40

И G3:



	precision	recall	f1-score	support
0	1.00	0.60	0.75	10
1	0.88	1.00	0.94	30
accuracy			0.90	40
macro avg	0.94	0.80	0.84	40
weighted avg	0.91	0.90	0.89	40

Следовательно, можно сделать вывод, что с помощью построенной модели можно делать прогноз G1, G2, G3.

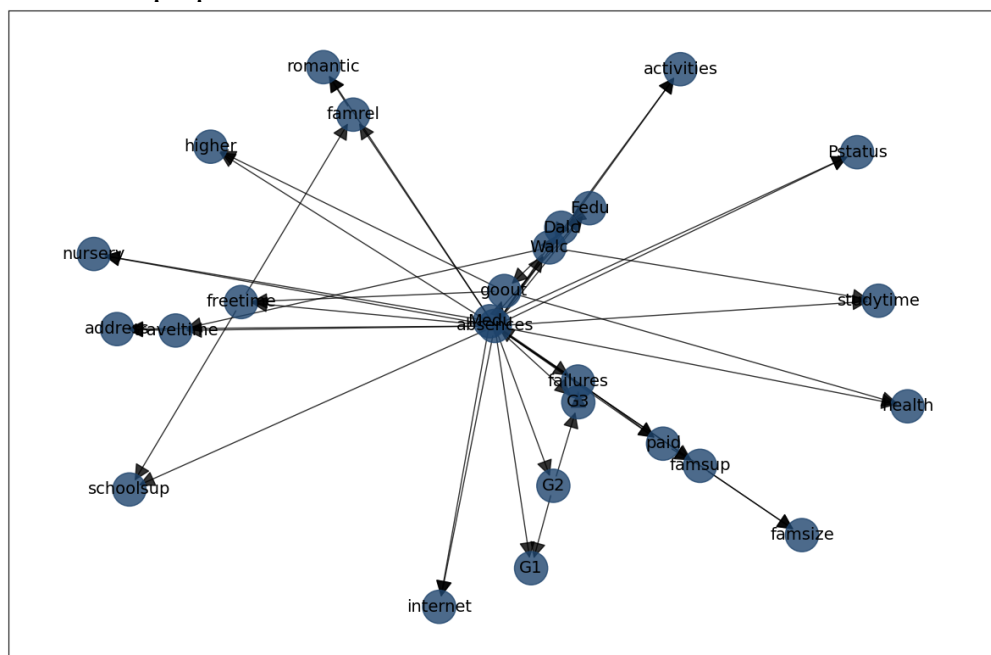
Tree-augmented Naive Bayes (TAN):

Аналогично для TAN.

В качестве class_node взят 'absences'. Вычислены оценки структуры ['k2', 'bds', 'bic', 'bdeu'] для сравнения моделей (чем выше, тем лучше). В результате для scoretype взят 'k2', т.к.:

```
'structure_scores': {'k2': -13082.169609157416,
'bds': -50533.69073791135,
'bic': -150483.7283041072,
'bdeu': -18545.96041896461}}
```

Полученный граф:



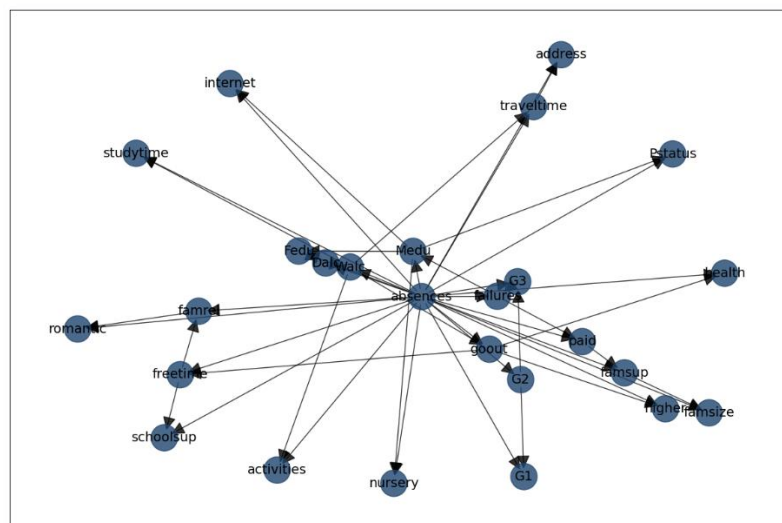
Пример CPD:

	absences	absences (0)	absences (1)	absences (2)
G2 (0)	0.497557003257329	0.4301639344262296	0.49085365853658536	
G2 (1)	0.502442996742671	0.5698360655737705	0.5091463414634146	

Оценка качества модели:

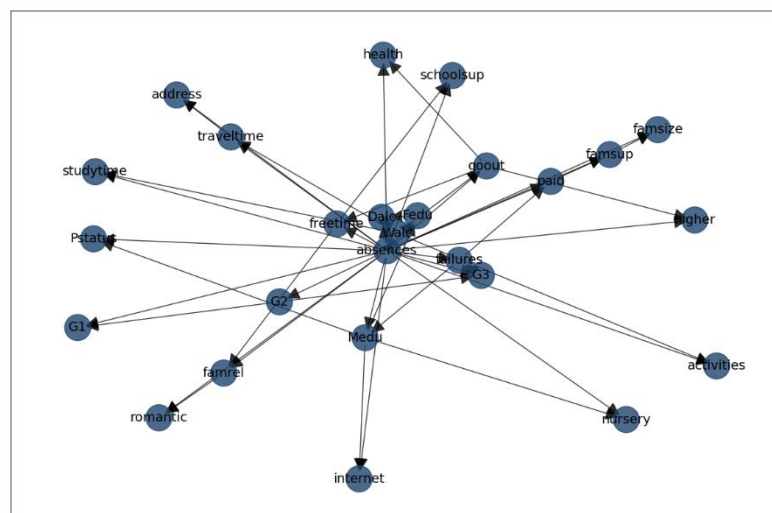
[[7 2] [2 29]]				
	precision	recall	f1-score	support
0	0.78	0.78	0.78	9
1	0.94	0.94	0.94	31
accuracy			0.90	40
macro avg	0.86	0.86	0.86	40
weighted avg	0.90	0.90	0.90	40

Для G2:



[[8 1] [0 31]]				
	precision	recall	f1-score	support
0	1.00	0.89	0.94	9
1	0.97	1.00	0.98	31
accuracy			0.97	40
macro avg	0.98	0.94	0.96	40
weighted avg	0.98	0.97	0.97	40

Для G3:



```
[[ 6 4]
 [ 0 30]]
```

	precision	recall	f1-score	support
0	1.00	0.60	0.75	10
1	0.88	1.00	0.94	30
accuracy			0.90	40
macro avg	0.94	0.80	0.84	40
weighted avg	0.91	0.90	0.89	40

Аналогично, можно сделать вывод, что с помощью построенной модели можно делать прогноз G1, G2, G3.

4. Исследование на данных *student-por.csv*.

Аналогично.

Исходные данные:

	school	sex	age	address	famsize	Pstatus	Medu	Fedu	Mjob	Fjob	...	famrel	freetime	goout	Dalc	Walc	health	absences	G1	G2	G3
0	GP	F	18	U	GT3	A	4	4	at_home	teacher	...	4	3	4	1	1	3	4	0	11	11
1	GP	F	17	U	GT3	T	1	1	at_home	other	...	5	3	3	1	1	3	2	9	11	11
2	GP	F	15	U	LE3	T	1	1	at_home	other	...	4	3	2	2	3	3	6	12	13	12
3	GP	F	15	U	GT3	T	4	2	health	services	...	3	2	2	1	1	5	0	14	14	14
4	GP	F	16	U	GT3	T	3	3	other	other	...	4	3	2	1	2	5	0	11	13	13
...
644	MS	F	19	R	GT3	T	2	3	services	other	...	5	4	2	1	2	5	4	10	11	10
645	MS	F	18	U	LE3	T	3	1	teacher	services	...	4	3	4	1	1	1	4	15	15	16
646	MS	F	18	U	GT3	T	1	1	other	other	...	1	1	1	1	1	5	6	11	12	9
647	MS	M	17	U	LE3	T	3	1	services	services	...	2	4	5	3	4	2	6	10	10	10
648	MS	M	18	R	LE3	T	3	2	services	other	...	4	4	1	3	4	5	4	10	11	11

Предобработанные данные:

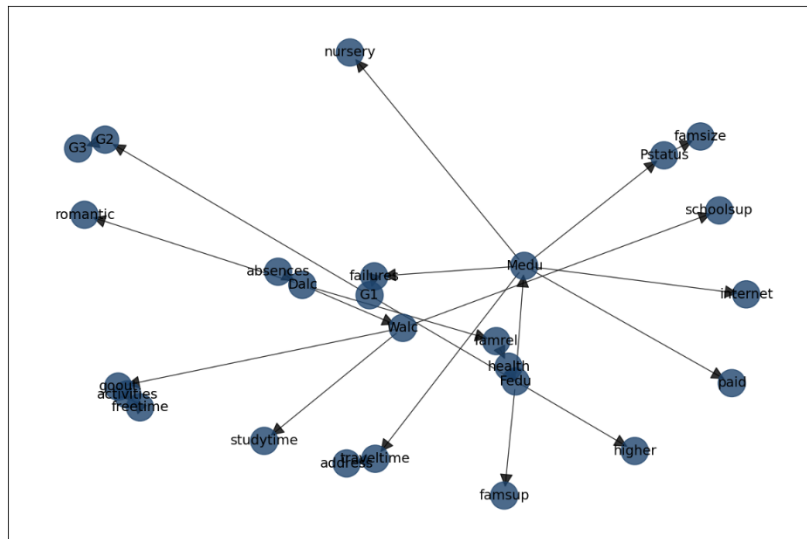
	address	famsize	Pstatus	Medu	Fedu	traveltime	studytime	failures	schoolsup	famsup	...	famrel	freetime	goout	Dalc	Walc	health	absences	G1	G2	G3
0	1	0	0	4	4	2	1	1	1	0	...	4	3	4	1	1	3	1	0	1	1
1	1	0	1	1	1	1	1	1	0	1	...	5	3	3	1	1	3	1	0	1	1
2	1	1	1	1	1	1	1	1	1	0	...	4	3	2	2	3	3	1	1	1	1
3	1	0	1	4	2	1	0	1	0	1	...	3	2	2	1	1	5	2	1	1	1
4	1	0	1	3	3	1	1	1	0	1	...	4	3	2	1	2	5	2	1	1	1
...
644	0	0	1	2	3	1	0	0	0	0	...	5	4	2	1	2	5	1	1	1	1
645	1	1	1	3	1	1	1	1	0	1	...	4	3	4	1	1	1	1	1	1	1
646	1	0	1	1	1	2	1	1	0	0	...	1	1	1	1	1	5	1	1	1	0
647	1	1	1	3	1	2	1	1	0	0	...	2	4	5	3	4	2	1	1	1	1
648	0	1	1	3	2	3	1	1	0	0	...	4	4	1	3	4	5	1	1	1	1

649 rows x 26 columns

Chow-liu

В качестве root_node ВЗЯТ 'absences'. Для scoretype ВЗЯТ 'k2', т.к.:
 'structure_scores': {'k2': -11688.757355927399,
 'bds': -11965.378046580681,
 'bic': -11959.818775040096,
 'bdeu': -11722.7765333847}}

Граф:



CPD of absences:

absences (0)	0.251473
absences (1)	0.400463
absences (2)	0.348064

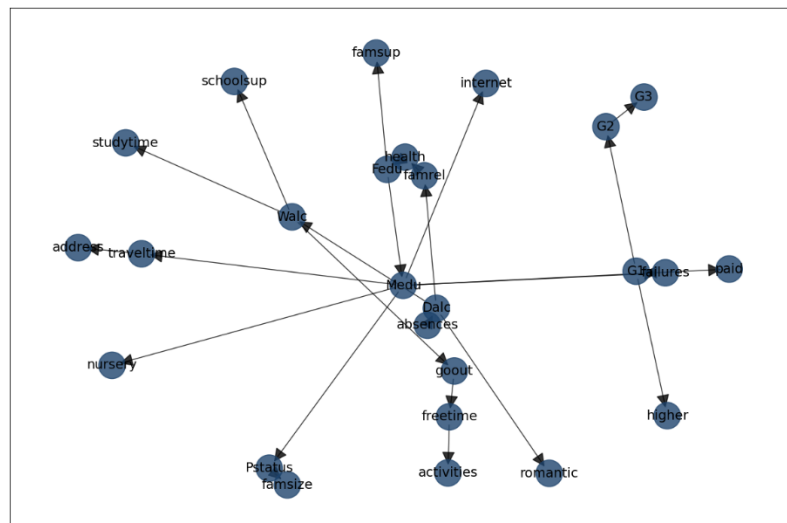
CPD of G1:

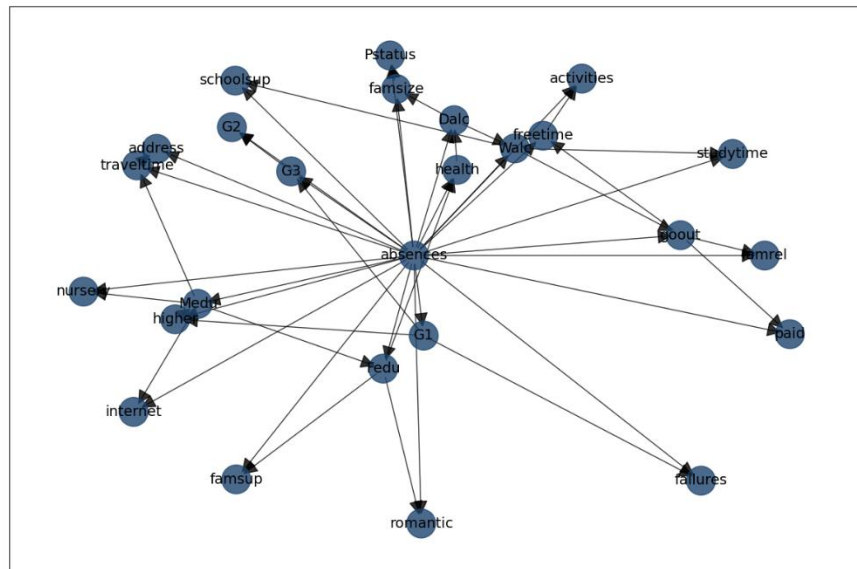
failures	failures (0)	failures (1)
G1 (0)	0.531986531986532	0.3323232323232323
G1 (1)	0.468013468013468	0.6676767676767676

[[5 7]
[0 53]]

	precision	recall	f1-score	support
0	1.00	0.42	0.59	12
1	0.88	1.00	0.94	53
accuracy			0.89	65
macro avg	0.94	0.71	0.76	65
weighted avg	0.90	0.89	0.87	65

G2:





```
[[ 7 2]
 [ 1 55]]
```

	precision	recall	f1-score	support
0	0.88	0.78	0.82	9
1	0.96	0.98	0.97	56
accuracy			0.95	65
macro avg	0.92	0.88	0.90	65
weighted avg	0.95	0.95	0.95	65

Как можно видеть из полученных результатов, построенные модели можно использовать для прогноза G1, G2, G3.

5. Линейная регрессия

Исходные данные student-mat.csv:

G1:

```
[[ 4 10]
 [ 4 22]]
```

	precision	recall	f1-score	support
0	0.50	0.29	0.36	14
1	0.69	0.85	0.76	26
accuracy			0.65	40
macro avg	0.59	0.57	0.56	40
weighted avg	0.62	0.65	0.62	40

G2:

```
[[ 6 7]
 [ 3 24]]
```

	precision	recall	f1-score	support
0	0.67	0.46	0.55	13
1	0.77	0.89	0.83	27
accuracy			0.75	40
macro avg	0.72	0.68	0.69	40
weighted avg	0.74	0.75	0.74	40

G3:

```
[[ 3 9]
 [ 3 25]]
      precision    recall  f1-score   support

     0       0.50      0.25      0.33        12
     1       0.74      0.89      0.81        28

 accuracy          0.70         40
 macro avg          0.62         40
 weighted avg       0.66         40
```

Как можно видеть, в случае линейной регрессии модель чаще ошибается при определении людей, которые не сдали. Для тех, кто сдал, модель работает хорошо.

Исходные данные student-por.csv:

G1:

```
[[ 6 6]
 [ 2 51]]
      precision    recall  f1-score   support

     0       0.75      0.50      0.60        12
     1       0.89      0.96      0.93        53

 accuracy          0.88         65
 macro avg          0.82         65
 weighted avg       0.87         65
```

G2:

```
[[ 7 7]
 [ 1 50]]
      precision    recall  f1-score   support

     0       0.88      0.50      0.64        14
     1       0.88      0.98      0.93        51

 accuracy          0.88         65
 macro avg          0.88         65
 weighted avg       0.88         65
```

G3:

```
[[ 0 9]
 [ 2 54]]
      precision    recall  f1-score   support

     0       0.00      0.00      0.00         9
     1       0.86      0.96      0.91        56

 accuracy          0.83         65
 macro avg          0.43         65
 weighted avg       0.74         65
```

Здесь аналогично, модель часто ошибается при определении людей, которые не сдали.

В целом, можно сделать вывод, что Байесовская сеть для исходных данных работает лучше, чем модель линейной регрессии.

Приложение 1. Код

```
# import libraries

! pip install causalnex
!pip install pgmpy
import numpy as np
import pandas as pd
from time import time
from sklearn.metrics import f1_score
import pgmpy.estimators as ests

import os
import random
import warnings
import numpy as np
import pandas as pd
import networkx as nx
import statsmodels
from scipy.io import arff
from scipy.signal._signaltools import _centered
import matplotlib.pyplot as plt
import pgmpy.estimators as ests
from pgmpy.models import BayesianNetwork
from pgmpy.metrics import structure_score
from sklearn.preprocessing import MinMaxScaler
from sklearn.metrics import mean_squared_error

# warnings.simplefilter("ignore")
!pip install bnlearn
import bnlearn as bn
from sklearn.preprocessing import LabelEncoder

# Learn the DAG using Chow-liu

model_cl = bn.structure_learning.fit(struct_data, methodtype='cl', root_node='absences', scoretype= 'k2')
# Plot detected DAG
G = bn.plot(model_cl)

# Learn the DAG using Tree-augmented Naive Bayes

model_tan = bn.structure_learning.fit(struct_data, methodtype='tan', class_node='absences', scoretype= 'k2')
# Plot detected DAG
G = bn.plot(model_tan)

# Read data

# student_mat_data = pd.read_csv("student-mat.csv")
```

```

# data = student_mat_data

student_por_data = pd.read_csv("student-por.csv")
data = student_por_data

# data preprocessing

data.notnull()

struct_data = data
discretised_data = struct_data.copy()

data_vals = {col: struct_data[col].unique() for col in struct_data.columns
}

failures_map = {v: 'no-failure' if v == [0]
                 else 'have-failure' for v in data_vals['failures']}
studytime_map = {v: 'short-studytime' if v in [1,2]
                 else 'long-studytime' for v in data_vals['studytime']}

discretised_data["failures"] = discretised_data["failures"].map(failures_map)
discretised_data["studytime"] = discretised_data["studytime"].map(studytime_map)

from causalnex.discretiser import Discretiser

discretised_data["absences"] = Discretiser(method="fixed",
                                           numeric_split_points=[1, 10]).transform(discretised_data["absences"].values)
discretised_data["G1"] = Discretiser(method="fixed",
                                     numeric_split_points=[10]).transform(discretised_data["G1"].values)
discretised_data["G2"] = Discretiser(method="fixed",
                                     numeric_split_points=[10]).transform(discretised_data["G2"].values)
discretised_data["G3"] = Discretiser(method="fixed",
                                     numeric_split_points=[10]).transform(discretised_data["G3"].values)

absences_map = {0: "No-absence", 1: "Low-absence", 2: "High-absence"}

G1_map = {0: "Fail", 1: "Pass"}
G2_map = {0: "Fail", 1: "Pass"}
G3_map = {0: "Fail", 1: "Pass"}

discretised_data["absences"] = discretised_data["absences"].map(absences_map)
discretised_data["G1"] = discretised_data["G1"].map(G1_map)
discretised_data["G2"] = discretised_data["G2"].map(G2_map)

```

```

discretised_data["G3"] = discretised_data["G3"].map(G3_map)

G_ = "G1"
struct_data = discretised_data.copy()
non_numeric_columns = list(struct_data.select_dtypes(exclude=[np.number]).
columns)

from sklearn.preprocessing import LabelEncoder
le = LabelEncoder()

for col in non_numeric_columns:
    struct_data[col] = le.fit_transform(struct_data[col])

drop_col = ['school', 'sex', 'age', 'Mjob', 'Fjob', 'reason', 'guardian']
struct_data = struct_data.drop(columns=drop_col)

from sklearn.model_selection import train_test_split
train, test = train_test_split(struct_data, train_size=0.9, test_size=0.1,
    random_state=7)
# Learn the DAG using Chow-liu

model_cl = bn.structure_learning.fit(train, methodtype='cl', root_node='absences', scoretype= 'k2')
# Plot detected DAG
G = bn.plot(model_cl)
# Parameter learning
model = bn.parameter_learning.fit(model_cl, train, verbose=3);

# Learn the DAG using Tree-augmented Naive Bayes

# model_tan = bn.structure_learning.fit(struct_data, methodtype='tan', class_node='absences', scoretype= 'k2')
# # Plot detected DAG
# G = bn.plot(model_tan)
# # Parameter learning
# model = bn.parameter_learning.fit(model_tan, train, verbose=3);

Pout = bn.predict(model, test, variables=[G_])
# print(Pout)

from sklearn.metrics import classification_report, confusion_matrix, accuracy_score
predictions = bn.predict(model, test, variables=[G_])[G_]
print("")
print(confusion_matrix(test[G_], predictions))
print(classification_report(test[G_], predictions))

# Linear Regression
from sklearn import linear_model

```

```

from sklearn.linear_model import LinearRegression
# Extract feature columns
feature_cols = list(struct_data.columns[:-3])
#feature_cols

# Extract target column 'passed'
#target_col = student_data.columns[-3:]
target_col_G1 = struct_data.columns[-3]
target_col_G2 = struct_data.columns[-2]
target_col_G3 = struct_data.columns[-1]

target_col = target_col_G3
#target_col

# Separate the data into feature data and target data (X_all and y_all, re
spectively)
X_all = struct_data[feature_cols]
y_all = struct_data[target_col]

# split the dataset
X_train, X_test, y_train, y_test = train_test_split(X_all, y_all, test_siz
e=0.1, random_state=0)

model = LinearRegression().fit(X_train, y_train)
y_pred = model.predict(X_test)

print("")
print(confusion_matrix(np.array(y_test), np rint(y_pred)))
print(classification_report(np.array(y_test), np rint(y_pred)))

```

Приложение 2. CPD.

```

[bnlearn] >CPD of G3:
+-----+-----+-----+-----+
| absences | absences(0)          | absences(1)          | absences(2)          |
+-----+-----+-----+-----+
| G3(0)    | 0.49022801302931596 | 0.4026229508196722 | 0.49085365853658536 |
+-----+-----+-----+-----+
| G3(1)    | 0.509771986970684  | 0.5973770491803279 | 0.5091463414634146  |
+-----+-----+-----+-----+
[bnlearn] >CPD of romantic:
+-----+-----+-----+-----+
--+
| absences | absences(0)          | absences(1)          | absences(2)          |
|
+-----+-----+-----+-----+
--+
| romantic(0) | 0.512214983713355   | 0.5836065573770491  |
0.5274390243902439 |
+-----+-----+-----+-----+
--+
| romantic(1) | 0.48778501628664495 | 0.41639344262295086 |
0.4725609756097561 |

```

```

+-----+-----+-----+-----+
--+
[bnlearn] >CPD of G2:
+-----+-----+-----+
| G3      | G3(0)                | G3(1)                |
+-----+-----+-----+
| G2(0)   | 0.5854838709677419   | 0.3727891156462585   |
+-----+-----+-----+
| G2(1)   | 0.41451612903225804 | 0.6272108843537415   |
+-----+-----+-----+
[bnlearn] >CPD of failures:
+-----+-----+-----+
| G3      | G3(0)                | G3(1)                |
+-----+-----+-----+
| failures(0) | 0.482258064516129   | 0.3795918367346939   |
+-----+-----+-----+
| failures(1) | 0.5177419354838709 | 0.6204081632653061   |
+-----+-----+-----+
[bnlearn] >CPD of G1:
+-----+-----+-----+
| G2      | G2(0)                | G2(1)                |
+-----+-----+-----+
| G1(0)   | 0.5667189952904239   | 0.3788300835654596   |
+-----+-----+-----+
| G1(1)   | 0.43328100470957615 | 0.6211699164345403   |
+-----+-----+-----+
[bnlearn] >CPD of Medu:
+-----+-----+-----+
| failures | failures(0)          | failures(1)          |
+-----+-----+-----+
| Medu(0)  | 0.17474048442906576 | 0.13127413127413126 |
+-----+-----+-----+
| Medu(1)  | 0.21107266435986158 | 0.16859716859716858 |
+-----+-----+-----+
| Medu(2)  | 0.21107266435986158 | 0.22651222651222652 |
+-----+-----+-----+
| Medu(3)  | 0.21453287197231835 | 0.21492921492921493 |
+-----+-----+-----+
| Medu(4)  | 0.18858131487889274 | 0.25868725868725867 |
+-----+-----+-----+
[bnlearn] >CPD of higher:
+-----+-----+-----+
| failures | failures(0)          | failures(1)          |
+-----+-----+-----+
| higher(0) | 0.45501730103806226 | 0.32947232947232946 |
+-----+-----+-----+
| higher(1) | 0.5449826989619377 | 0.6705276705276705 |
+-----+-----+-----+
[bnlearn] >CPD of Fedu:
+-----+-----+-----+
| Medu     | Medu(0)              | ... | Medu(4)              |
+-----+-----+-----+
| Fedu(0)  | 0.19704433497536947 | ... | 0.13225806451612904 |
+-----+-----+-----+
| Fedu(1)  | 0.2019704433497537  | ... | 0.13225806451612904 |
+-----+-----+-----+
| Fedu(2)  | 0.20689655172413793 | ... | 0.1774193548387097   |
+-----+-----+-----+
| Fedu(3)  | 0.19704433497536947 | ... | 0.23870967741935484 |
+-----+-----+-----+
| Fedu(4)  | 0.19704433497536947 | ... | 0.3193548387096774   |
+-----+-----+-----+
[bnlearn] >CPD of internet:
+-----+-----+-----+

```



```

| Medu          | Medu(0)          | ... | Medu(4)          |
+-----+-----+-----+-----+
| internet(0)   | 0.4975369458128079 | ... | 0.34838709677419355 |
+-----+-----+-----+-----+
| internet(1)   | 0.5024630541871922 | ... | 0.6516129032258065 |
+-----+-----+-----+-----+
[bnlearn] >CPD of nursery:
+-----+-----+-----+-----+
| Medu          | Medu(0)          | ... | Medu(4)          |
+-----+-----+-----+-----+
| nursery(0)    | 0.5024630541871922 | ... | 0.36129032258064514 |
+-----+-----+-----+-----+
| nursery(1)    | 0.4975369458128079 | ... | 0.6387096774193548 |
+-----+-----+-----+-----+
[bnlearn] >CPD of Pstatus:
+-----+-----+-----+-----+
| Medu          | Medu(0)          | ... | Medu(4)          |
+-----+-----+-----+-----+
| Pstatus(0)    | 0.49261083743842365 | ... | 0.3548387096774194 |
+-----+-----+-----+-----+
| Pstatus(1)    | 0.5073891625615764 | ... | 0.6451612903225806 |
+-----+-----+-----+-----+
[bnlearn] >CPD of activities:
+-----+-----+-----+
| Medu          | ... | Medu(4)          |
+-----+-----+-----+
| activities(0)  | ... | 0.47096774193548385 |
+-----+-----+-----+
| activities(1)  | ... | 0.5290322580645161 |
+-----+-----+-----+
[bnlearn] >CPD of paid:
+-----+-----+-----+
| higher        | higher(0)          | higher(1)          |
+-----+-----+-----+
| paid(0)       | 0.5163776493256262 | 0.5047846889952153 |
+-----+-----+-----+
| paid(1)       | 0.4836223506743738 | 0.49521531100478466 |
+-----+-----+-----+
[bnlearn] >CPD of traveltime:
+-----+-----+-----+
| Fedu          | ... | Fedu(4)          |
+-----+-----+-----+
| traveltime(1) | ... | 0.3807829181494662 |
+-----+-----+-----+
| traveltime(2) | ... | 0.24555160142348753 |
+-----+-----+-----+
| traveltime(3) | ... | 0.18861209964412812 |
+-----+-----+-----+
| traveltime(4) | ... | 0.18505338078291814 |
+-----+-----+-----+
[bnlearn] >CPD of famsup:
+-----+-----+-----+
| paid          | paid(0)          | paid(1)          |
+-----+-----+-----+
| famsup(0)     | 0.5057971014492754 | 0.4345864661654135 |
+-----+-----+-----+
| famsup(1)     | 0.49420289855072463 | 0.5654135338345865 |
+-----+-----+-----+
[bnlearn] >CPD of address:
+-----+-----+-----+
| traveltime    | traveltime(1)      | ... | traveltime(4)      |
+-----+-----+-----+
| address(0)    | 0.32426778242677823 | ... | 0.5038759689922481 |
+-----+-----+-----+

```

```

| address(1) | 0.6757322175732218 | ... | 0.49612403100775193 |
+-----+-----+-----+
[bnlearn] >CPD of Walc:
+-----+-----+-----+
| traveltime | traveltime(1) | ... | traveltime(4) |
+-----+-----+-----+
| Walc(1) | 0.28870292887029286 | ... | 0.19767441860465115 |
+-----+-----+-----+
| Walc(2) | 0.20711297071129708 | ... | 0.19767441860465115 |
+-----+-----+-----+
| Walc(3) | 0.20711297071129708 | ... | 0.1937984496124031 |
+-----+-----+-----+
| Walc(4) | 0.1589958158995816 | ... | 0.1937984496124031 |
+-----+-----+-----+
| Walc(5) | 0.13807531380753138 | ... | 0.21705426356589147 |
+-----+-----+-----+
[bnlearn] >CPD of Dalc:
+-----+-----+-----+
| Walc | Walc(1) | ... | Walc(5) |
+-----+-----+-----+
| Dalc(1) | 0.5194029850746269 | ... | 0.1894273127753304 |
+-----+-----+-----+
| Dalc(2) | 0.12238805970149254 | ... | 0.19823788546255505 |
+-----+-----+-----+
| Dalc(3) | 0.11940298507462686 | ... | 0.2026431718061674 |
+-----+-----+-----+
| Dalc(4) | 0.11940298507462686 | ... | 0.19383259911894274 |
+-----+-----+-----+
| Dalc(5) | 0.11940298507462686 | ... | 0.21585903083700442 |
+-----+-----+-----+
[bnlearn] >CPD of goout:
+-----+-----+-----+
| Walc | Walc(1) | ... | Walc(5) |
+-----+-----+-----+
| goout(1) | 0.16417910447761194 | ... | 0.1762114537444934 |
+-----+-----+-----+
| goout(2) | 0.25970149253731345 | ... | 0.18502202643171806 |
+-----+-----+-----+
| goout(3) | 0.24776119402985075 | ... | 0.19383259911894274 |
+-----+-----+-----+
| goout(4) | 0.191044776119403 | ... | 0.2026431718061674 |
+-----+-----+-----+
| goout(5) | 0.1373134328358209 | ... | 0.2422907488986784 |
+-----+-----+-----+
[bnlearn] >CPD of studytime:
+-----+-----+-----+
| Walc | ... | Walc(5) |
+-----+-----+-----+
| studytime(0) | ... | 0.45374449339207046 |
+-----+-----+-----+
| studytime(1) | ... | 0.5462555066079295 |
+-----+-----+-----+
[bnlearn] >CPD of schoolsup:
+-----+-----+-----+
| Walc | ... | Walc(5) |
+-----+-----+-----+
| schoolsup(0) | ... | 0.5462555066079295 |
+-----+-----+-----+
| schoolsup(1) | ... | 0.45374449339207046 |
+-----+-----+-----+
[bnlearn] >CPD of health:
+-----+-----+-----+
| Dalc | Dalc(1) | ... | Dalc(5) |
+-----+-----+-----+

```

```

| health(1) | 0.1547085201793722 | ... | 0.19617224880382775 |
+-----+-----+-----+-----+
| health(2) | 0.15022421524663676 | ... | 0.19138755980861244 |
+-----+-----+-----+-----+
| health(3) | 0.24439461883408073 | ... | 0.19617224880382775 |
+-----+-----+-----+-----+
| health(4) | 0.18385650224215247 | ... | 0.21052631578947367 |
+-----+-----+-----+-----+
| health(5) | 0.26681614349775784 | ... | 0.20574162679425836 |
+-----+-----+-----+-----+
[bnlearn] >CPD of famsize:
+-----+-----+-----+-----+
| Dalc      | Dalc(1)          | ... | Dalc(5)          |
+-----+-----+-----+-----+
| famsize(0) | 0.6524663677130045 | ... | 0.5167464114832536 |
+-----+-----+-----+-----+
| famsize(1) | 0.3475336322869955 | ... | 0.48325358851674644 |
+-----+-----+-----+-----+
[bnlearn] >CPD of freetime:
+-----+-----+-----+-----+
| goout     | goout(1)         | ... | goout(5)         |
+-----+-----+-----+-----+
| freetime(1) | 0.1981981981981982 | ... | 0.1659919028340081 |
+-----+-----+-----+-----+
| freetime(2) | 0.20270270270270271 | ... | 0.17408906882591094 |
+-----+-----+-----+-----+
| freetime(3) | 0.20270270270270271 | ... | 0.21862348178137653 |
+-----+-----+-----+-----+
| freetime(4) | 0.20270270270270271 | ... | 0.21862348178137653 |
+-----+-----+-----+-----+
| freetime(5) | 0.19369369369369369 | ... | 0.22267206477732793 |
+-----+-----+-----+-----+
[bnlearn] >CPD of famrel:
+-----+-----+-----+-----+
| health    | health(1)        | ... | health(5)        |
+-----+-----+-----+-----+
| famrel(1) | 0.17842323651452283 | ... | 0.13069908814589665 |
+-----+-----+-----+-----+
| famrel(2) | 0.17427385892116182 | ... | 0.1337386018237082 |
+-----+-----+-----+-----+
| famrel(3) | 0.1908713692946058 | ... | 0.1762917933130699 |
+-----+-----+-----+-----+
| famrel(4) | 0.23651452282157676 | ... | 0.3252279635258359 |
+-----+-----+-----+-----+
| famrel(5) | 0.21991701244813278 | ... | 0.23404255319148937 |
+-----+-----+-----+-----+

```