



ЦЕНТР
ДОПОЛНИТЕЛЬНОГО
ОБРАЗОВАНИЯ
МГТУ им. Н.Э. Баумана

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА

на тему:

«Прогнозирование конечных свойств новых материалов
(композиционных материалов)».

Старостина Ярослава Константиновна



Постановка задачи. Разведывательный анализ данных.

Целью данной работы является прогнозирование конечных свойств новых материалов (композиционных материалов).

Кейс основан на реальных производственных задачах Центра НТИ «Цифровое материаловедение: новые материалы и вещества» (структурное подразделение МГТУ им. Н.Э. Баумана).

Для исследовательской работы были даны 2 файла: X_br.xlsx, состоящий из 1023 строки и 11 столбцов, и X_pur.xlsx, состоящий из 1040 строк и 4 столбцов. Объединение делать по индексу тип объединения INNER, как указано в задании.

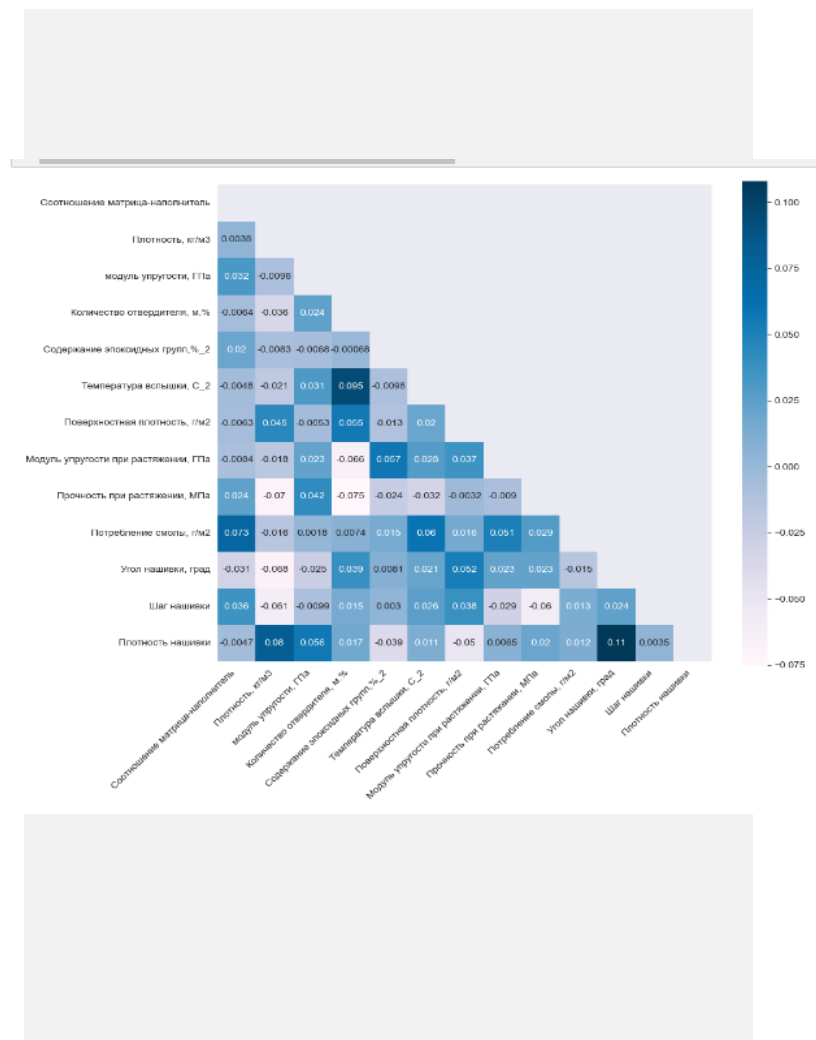
Все переменные содержат значения float64, качественные характеристики отсутствуют. Пропусков не имеется. Ни одна из записей не является NaN, очистка не требуется. Для удаления выбросов был применен метод межквартильных расстояний, далее была проведена замены пустых значений на медиану.



Постановка задачи. Разведывательный анализ данных.

Данные не имеют чётко выраженной зависимости, что подтверждает тепловая карта с матрицей корреляции и матрицы диаграмм рассеяния.

В соответствии с теорией композитных материалов на качество материала влияет температура вспышки и количество отвердителя из-за взаимодействия отвердителя с матрицей и наполнителем под влиянием температуры. Угол нашивки и плотность нашивки, несомненно, оказывают влияние на свойства материала, что подтверждается максимальным значением 0.11. Но в целом корреляция между всеми параметрами очень близка к 0, корреляционные связи между переменными не наблюдаются.





Предобработка данных. Разработка и обучение модели

Алгоритмы машинного обучения, как правило, работают лучше или быстрее, когда различные функции (переменные) имеют меньший масштаб. Поэтому перед обучением на них моделей машинного обучения данные обычно нормализуются.

В качестве параметра оценки выбран коэффициент детерминации (R^2); оптимизация подбора гиперпараметров модели с помощью выбора по сетке и перекрёстной проверки; подстановка оптимальных гиперпараметров в модель; оценка полученных данных.

Разработка и обучение моделей машинного обучения осуществлялась для двух выходных параметров: «Прочность при растяжении» и «Модуль упругости при растяжении» отдельно. Для решения применим

Модель после настройки гиперпараметров показала результат немного лучше. Однако, ниже, чем базовая модель. Все использованные модели не справились с задачей. Результат неудовлетворительный.



Предобработка данных. Разработка и обучение модели

```
jupyter Starostina Ya.... Последняя контрольная точка: час назад (автосохранение)
```

```
File Edit View Insert Cell Kernel Widgets Help Не доверять Pyt
```

```
r2_score_l1,  
r2_score_br,  
r2_score_pr,  
r2_score_mlp}}  
mae_r2_df = pd.DataFrame.from_dict(mae_r2_df, orient='index')  
mae_r2_df.T
```

Out[50]:

	Перепсеп	MAE	R2 score
0	Linear Regression	2.519962	-0.02778
1	SVR rbf	2.61258	-0.131368
2	SVR poly	2.933201	-0.490473
3	K Neighbors Regressor	2.491875	-0.008515
4	Decision Tree Regressor	2.522181	-0.027262
5	Random Forest Regressor	2.564905	-0.056906
6	Gradient Boosting Regressor	2.57924	-0.11011
7	Stochastic Gradient Descent Regressor	2.519193	-0.026995
8	Lasso Lars	2.509005	-0.021567
9	Bayesian Ridge	2.505908	-0.019647
10	Poisson Regressor	2.519433	-0.02737
11	Multi-layer Perceptron regressor	3.217804	-0.747803

```
jupyter Starostina Ya.K. KOMPOSIT_4 (автосохранение)
```

```
File Edit View Insert Cell Kernel Widgets Help Не доверять
```

```
mae_r2_df = pd.DataFrame.from_dict(mae_r2_df, orient='index')  
mae_r2_df.T
```

Out[57]:

	Перепсеп	MAE	R2 score
0	Linear Regression	373.756781	-0.000892
1	SVR rbf	379.096126	-0.022535
2	SVR poly	385.073723	-0.075725
3	K Neighbors Regressor	370.785139	0.005681
4	Decision Tree Regressor	371.652748	0.002649
5	Random Forest Regressor	373.208372	-0.002043
6	Gradient Boosting Regressor	381.594145	-0.044607
7	Stochastic Gradient Descent Regressor	373.390259	-0.000271
8	Lasso Lars	373.756781	-0.000892
9	Bayesian Ridge	372.165846	-0.000703
10	Poisson Regressor	373.838099	-0.001265
11	Multi-layer Perceptron regressor	425.335798	-0.253642

Результаты оценки точности по MAE и R2 для прогнозирования величины "Модуль упругости при растяжении, ГПа" и "Прочность при растяжении, МПа" соответственно.



Нейронная сеть, для рекомендации соотношения «матрица – наполнитель».

```
jupyter Starostina Ya.K. KOMPOSIT_5 (автосохранение)
File Edit View Insert Cell Kernel Widgets Help Не доверять Python 3 (ipykernel)
Ввод [4]: # Определим лучшие параметры при помощи обёртки для API Scikit-Learn
# https://ru-keras.com/scikit-learn-api/
def build_model(activation='tanh', optimizer='SGD'):
    my_model = Sequential()
    my_model.add(Normalization(axis=-1))
    my_model.add(Dense(units = 14 , activation= 'tanh'))
    my_model.add(Dense(units = 14 , activation= 'tanh'))
    my_model.add(Dropout(0.1))
    my_model.add(Dense(units = 1 , activation= 'softmax'))
    my_model.add(Normalization(axis=-1, invert=True))

    my_model.compile(loss='mean_squared_error',
                    optimizer=optimizer,
                    metrics=['mae', 'accuracy'])

    return my_model

my_model_kr = KerasRegressor(build_fn=build_model,
                             epochs=10,
                             batch_size=5,
                             verbose=0)

parameters = {'activation': ['softmax', 'softplus', 'softsign', 'relu', 'tanh',
                             'batch_size': [4, 8, 12, 16, 20],
                             'epochs': [10, 50, 100, 200, 300],
                             'optimizer': ['SGD', 'RMSprop', 'Adagrad', 'Adadelta', 'Adam', 'N

grid = GridSearchCV(estimator=my_model_kr,
                    param_grid=parameters,
                    cv=10,
                    verbose=1, n_jobs=-1)

grid_result = grid.fit(x_train, y_train)

print("Best score: %f using %s" % (grid_result.best_score_, grid_result.best_par

Fitting 10 folds for each of 1200 candidates, totalling 12000 fits
Best score: -4.457734 using {'activation': 'softmax', 'batch_size': 12, 'epoch
s': 10, 'optimizer': 'SGD'}
```

```
jupyter Starostina Ya.K. KOMPOSIT_6 (автосохранение)
File Edit View Insert Cell Kernel Widgets Help Доверенный Python 3 (ipykernel)
Ввод [6]: y_pred_model = model.predict(x_test)
mae = mean_absolute_error(y_pred_model, y_test)
r2_score = r2_score(y_pred_model, y_test)
print('mean absolute error:', round(mae, 2))
print("r2 score: ", round(r2_score, 2))

10/10 [=====] - 0s 2ms/step
mean absolute error: 1.95
r2 score: 0.0

Ввод [7]: model.summary()

Model: "sequential"

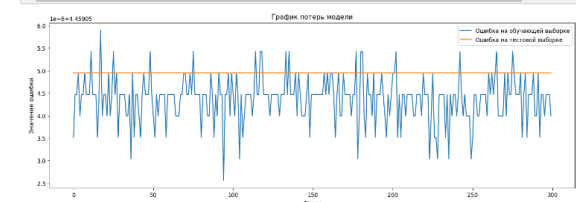
Layer (type) Output Shape Param #
=====
normalization (Normalizatio (None, 14) 29
n)
dense (Dense) (None, 14) 210
dense_1 (Dense) (None, 14) 210
dropout (Dropout) (None, 14) 0
dense_2 (Dense) (None, 1) 15
normalization_1 (Normalizat (None, 1) 3
ion)

Total params: 467
Trainable params: 435
Non-trainable params: 32
```

```
jupyter Starostina Ya... Последняя контрольная точка: 18.04.2023 (автосохранение)
File Edit View Insert Cell Kernel Widgets Help Доверенный Python 3 (ipykernel)
Ввод [5]: plt.figure(figsize = (17,5))
plt.plot(history.history['loss'], label = 'ошибка на обучающей выборке')
plt.plot(history.history['val_loss'], label = 'ошибка на тестовой выборке')
plt.title('График потерь модели')
plt.xlabel('Эпохи')
plt.ylabel('Значение ошибки')
plt.legend(['ошибка на обучающей выборке', 'ошибка на тестовой выборке'], loc=
plt.show()

Ввод [6]: y_pred_model = model.predict(x_test)
mae = mean_absolute_error(y_pred_model, y_test)
r2_score = r2_score(y_pred_model, y_test)
print('mean absolute error:', round(mae, 2))
print("r2 score: ", round(r2_score, 2))

10/10 [=====] - 0s 2ms/step
mean absolute error: 1.95
r2 score: 0.0
```



Построение нейронной сети после определения лучших параметров

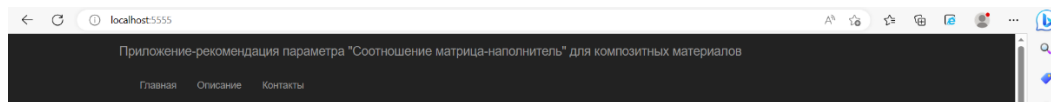
В результате можно сделать вывод о том, что нейронная сеть обучилась выдавать среднее значение соотношения «матрица-наполнитель».



ЦЕНТР
ДОПОЛНИТЕЛЬНОГО
ОБРАЗОВАНИЯ
МГТУ им. Н.Э. Баумана

Практическая часть

Разработка приложения



Приложение-рекомендация параметра "Соотношение матрица-наполнитель" для композитных материалов

Главная Описание Контакты

Плотность, кг/м3

Модуль упругости, ГПа

Количество отвердителя, м. %

Содержание эпоксидных групп, %_2

Температура вспышки, С_2

Поверхностная плотность, г/м2

Модуль упругости при растяжении, ГПа

Прочность при растяжении, МПа

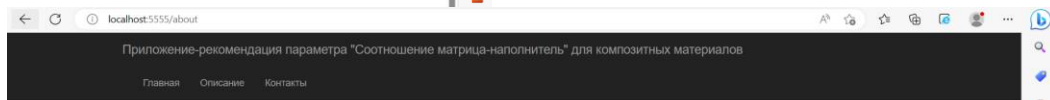
Потребление смолы, г/м2

Угол нашивки, град

Шаг нашивки

Плотность нашивки

© 2023 - Приложение выполнила: слушатель курса Старостина Ярослава Константиновна



Приложение-рекомендация параметра "Соотношение матрица-наполнитель" для композитных материалов

Главная Описание Контакты

Описание приложения.

Тема: Прогнозирование конечных свойств новых материалов (композиционных материалов).

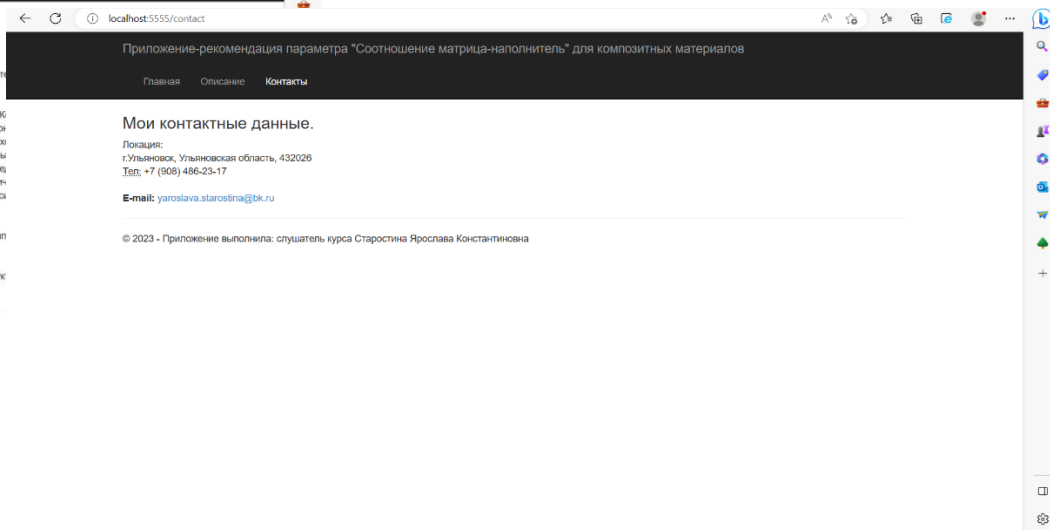
Актуальность: Созданные прогнозные модели помогут сократить количество проводимых испытаний, а также пополнить базу данных материалов, и цифровыми двойниками новых композитов.

Композиционные материалы - это искусственно созданные материалы, состоящие из нескольких других с четкой границей между ними. К свойствам, которые не наблюдаются у компонентов по отдельности. При этом композиты являются монолитным материалом, т.е. композит от друга без разрушения конструкции в целом. Яркий пример композита - железобетон. Бетон прекрасно сопротивляется сжатию, но плохо внутри бетона компенсирует его неспособность сопротивляться сжатию, формируя тем самым трещины, уникальные свойства. Современны другие материалы: полимеры, керамика, стеклянные и углеродные волокна, но данный принцип сохраняется. У такого подхода есть и не характеристики исходных компонентов, определить характеристики композита, состоящего из этих компонентов, достаточно проблематично. Есть два пути: физические испытания образцов материалов, или прогнозирование характеристик. Суть прогнозирования заключается в анализе элемента объема композита, на основе данных о характеристиках входящих компонентов (связующего и армирующего компонента).

На входе имеются данные о начальных свойствах компонентов композиционных материалов (количество связующего, наполнителя, темп т.д.). На выходе необходимо спрогнозировать параметр "Соотношение матрица-наполнитель".

Кейс основан на реальных производственных задачах Центра НТИ «Цифровое материаловедение: новые материалы и вещества» (структура Н.Э. Баумана).

© 2023 - Приложение выполнила: слушатель курса Старостина Ярослава Константиновна



Приложение-рекомендация параметра "Соотношение матрица-наполнитель" для композитных материалов

Главная Описание Контакты

Мои контактные данные.

Локация:
г.Ульяновск, Ульяновская область, 432026
Тел: +7 (908) 486-23-17
E-mail: yaroslava.starostina@bk.ru

© 2023 - Приложение выполнила: слушатель курса Старостина Ярослава Константиновна

Скриншоты веб-приложения



ЦЕНТР
ДОПОЛНИТЕЛЬНОГО
ОБРАЗОВАНИЯ
МГТУ им. Н.Э. Баумана



do.bmstu.ru