

PdM Mid-Term 智能性维护期中测试

Student Name / 姓名:

* Required

1. Questions : Why call Manufacturing 4.0 Smart? 智能制造的三大优势 Choose 3 * 5 points

Check all that apply.

- ☐ Efficiency – high productivity 高效
- ☐ Reliability – Reduce downtime 可靠
- ☐ Cost Effective – lower cost 低成本效应
- ☐ Low Initial Investment Needed 低投资

2. Questions : Which of the following are Smart Manufacturing Use Cases 下面哪些是智能制造的应用场景. Choose 选择 3 个 * 5 points

Check all that apply.

- ☐ Using Reinforcement Learning in AI for product design 智慧产品研发
- ☐ Process parameters optimization with AI 工艺参数的优化
- ☐ To estimate when maintenance should be performed, with AI. 预测性维护
- ☐ Use Social Media for Technical Support. 利用社交媒体做技术支持

3. Question . What is Additive Manufacturing * 5 points
什么是增材制造

4. Question : List Types of Sensors You Have Knowledge With *

5 points

列出你所知的工业传感器类型

5. Question . Maintenance. List Each Type of Maintenance's Pros and Cons. * 10 points

These four types of maintenances are: Reactive, Preventive, Condition Based and Predictive.

列出4类维护维修各自的优缺点。 分别是： 应对维修， 预防维修， 基于设备状态安排维护， 智能性维护

6. Question . Github, Gitbash. Which Of the following Is NOT a valid git-bash command? * 5 points

下面哪一个命令不是 git bash 命令?

Mark only one oval.

- ☐ git pull -a
- ☐ git stash
- ☐ git branch -a
- ☐ git remove all

7. Question . Get Familiar with Our Data. The following are our dataset. 熟悉我们 * 5 points
的项目所用的数据。

Find out which engine in the training dataset has the longest cycle? 在我们的训练数据集里，那个引擎起降落次数最高？如果不能运行，请写出 python 语句

Metadata about C-MAPSS / C-MAPSS 元数据

Data Set / 数据集	Train trajectories / 训练集	Test trajectories / 训练集	Conditions / 运行状态	Fault modes / 故障模式
FD001	100	100	One(sea level)	One(HPD Degradation / 退化)
FD002	260	259	Six	One(HPC degradation / 退化)
FD003	100	100	One	Two(HPC degradation, fan degradation / 风扇退化)
FD004	249	248	Six	Two(HPC degradation, fan degradation / 风扇退化)

8. Question . Read the following script.

* 5 points

阅读下面的程序：

```
def unzip_files(self, zip_file_name = None, remove_zipped = False):
    if not zip_file_name:
        zip_file_name = f'{self.FILE_PATHS["zip_data_path"]}/CMAPSS.zip'

    if not os.path.exists(self.FILE_PATHS["zip_data_path"]):
        os.makedirs(self.FILE_PATHS["zip_data_path"])

    with zipfile.ZipFile(zip_file_name, 'r') as zip_ref:
        zip_ref.extractall(self.FILE_PATHS["unzip_to_path"])

    if remove_zipped:
        os.remove(zip_file_name)
```

What does this function do? 这段程序起什么作用？

Mark only one oval.

- ☐ It zip the CMAPSS file 压缩CMAPSS 文件
- ☐ It unzip CMAPSS file or other zip file if provided as input 解压缩 CMAPSS 或输入的文件
- ☐ It does nothing but remove the zipfile 只是去掉压缩文件
- ☐ It does nothing but create a folder 只是创建一个文件夹

9. Question . Read the following code snippets: 阅读有效程序:

* 5 points

```
#df_train is DataFrame from self.read_data_files
df_train = self.read_data_files( file_name_str = "train", use_pd = use_pd,
sep = " ", columns = columns)

df_test = self.read_data_files( file_name_str = "test", use_pd = use_pd, sep
= " ", columns = columns)

df_train.iloc[:, [0,1]] = df_train.iloc[:, [0,1]].astype(int)
df_test.iloc[:, [0,1]] = df_test.iloc[:, [0,1]].astype(int)
```

What are changed with DataFrame df_train and df_test? df_train 和 df_test DataFrame 有什么变化?

Mark only one oval.

- ☐ It changes Nothing 没有变化
- ☐ It changes first two rows of df_train and df_test 把前面两行改变
- ☐ It changes first two columns of both df_train and df_test to 0, 1 把前面两列改变, 赋值0, 1
- ☐ It changes first two columns of both df_train and df_test to data type integer 把前面两列改变, 只是改变数据类型, 改成整数

10. Question . EDA. For conventional EDA, we may carry over EDA from the following perspectives. Choose 3

* 5 points

选择常用的从那些方面进行探索性分析。选三个

Check all that apply.

- ☐ Structures of Data 结构
- ☐ Which team to handle Data 那个小组处理数据
- ☐ Quality of Data 数据质量
- ☐ Content of Data 数据内容

11. Question . EDA on Data Content. Examine the following code snippet. * 5 points

What are the expected results.

DF_TRAIN is the training Dataset in form of DataFrame

阅读以下程序。预期的结果是什么？DF_TRAIN 是我们的训练数据集

```
tmp = DF_TRAIN.copy()
tmp.hist(bins=25, figsize=(15, 25), layout=(-1, 5), edgecolor="black")
plt.tight_layout();
```

12. Question . Feature Engineering. Transformation and Scaling is very * 5 points

common in Feature Engineering.

Explain Standard Scaler. (you may ref

to [sklearn.preprocessing.StandardScaler](#)) Math formula is highly

recommended. 特征工程。转换和缩放在特征工程中非常常见。解释标准定

标器。（你可以参考 sklearn.preprocessing.StandardScaler）强烈推荐引

用数学公式

13. Question . Feature Engineering. Transformation and Scaling is very common in Feature Engineering.

* 5 points

Explain Box-Cox transformation. Math formula is highly recommended. 问题 13. 特征工程。转换和缩放在特征工程中非常常见。解释 Box-Cox 变换。强烈推荐数学公式。

14. Question . Feature Selection.

* 5 points

for the following code snippet, here DF_TRAIN is our training dataset. And it has data of all settings (i.e., all FD001, FD002, FD003, FD004 are in the dataset).

问题 14. 特征选择。

对于以下代码片段，这里的 DF_TRAIN 是我们的训练数据集。它具有所有设置的数据（即所有 FD001、FD002、FD003、FD004 都在数据集中）。

```
tmp = DF_TRAIN.copy()
tmp["rul"] = tmp.groupby(["id", "Flag"])["cycle"].transform("max") -
tmp["cycle"]
qry_str = f"Flag=='{fault_op}"
tmp = tmp[tmp.eval(qry_str)]
tmp["Distance_To_Fail"] = pd.cut(tmp["rul"], bins=[0,20,50, 100, 200, 600],
include_lowest=True,right=True, labels=["Low", "Mid-Low", "Mid", "Mid-
High", "High"])
```

if we want to find out whether there is any relationship between Distance_To_Fail and Setting (i.e., FD00x), what stats test works the best? 如果我们想知道 Distance_To_Fail 和 Setting（即 FD00x）之间是否有任何关系，什么统计测试效果最好？

Mark only one oval.

☐ chi2

☐ Anova

☐ t-test

☐ Kendall rank

☐ Other: _____

15. Question . For first round of feature selection, which of the following action is Not recommended: 问题 16. 对于第一轮特征选择，不推荐以下哪个动作: * 5 points

Mark only one oval.

- ☐ Drop or handle Nulls 处理空值
- ☐ Remove constant or quasi constant features 移除常数或准常数特征
- ☐ Dispersion ratio / Laplacian Score 分散比/拉普拉斯分数
- ☐ Drop datetime columns 删除日期时间列
- ☐ Other: _____

16. Question . Feature Selection Methodologies. * 5 points
- Choose the conventional feature selection methodologies, choose 3
特征选择方法。
选择常规的特征选择方法，选择3

Mark only one oval.

- ☐ Filtering Method 过滤
- ☐ Wrapping Method 包装
- ☐ Deep Learning 深度学习
- ☐ Embedding Method 内嵌
- ☐ Other: _____

17. Question . Feature Selection Methodologies. For univariate feature selection, SelectKBest is a common process. Sklearn has the following implementation: * 5 points
- 特征选择方法。对于单一特征选择，SelectKBest 是一个常见的过程。Sklearn 有以下实现：

sklearn.feature_selection.SelectKBest(score_func=, *, k=10)

Which score_func can be used? Choose all apply
可以使用哪个 score_func?

Check all that apply.

- ☐ f_classif
☐ mutual_info_classif
☐ chi2
☐ f_regression
☐ Other: _____

18. Question . Feature Selection Methodologies. Why LASSO can be used as * 10 points
an effective way for feature selection?

You may refer to: menu_C_b_FS_Biligual.ipynb

特征选择方法。为什么 LASSO 可以作为特征选择的有效方法？
你可以参考 menu_C_b_FS_Biligual
