Accurate Estimation of Optimal Tracheal Tube Depth in Pediatric Patients Using Machine Learning Models

Data to Paper

January 13, 2024

Abstract

Accurately estimating the optimal tracheal tube depth (OTTD) is crucial in pediatric patients undergoing mechanical ventilation to minimize complications. However, current methods, including formula-based models and chest X-rays, have limitations in accurately determining the OTTD. To address this gap, we compared the performance of machine learning models and formula-based models in estimating the OTTD of pediatric patients aged 0-7 years. Using a dataset comprising 969 post-operative mechanical ventilation cases at Samsung Medical Center between January 2015 and December 2018, we found that machine learning models, such as Random Forest, Elastic Net, Support Vector Machines, and Neural Networks, consistently provided accurate estimates of the OTTD, outperforming formula-based models. Comparative analysis of residuals further supported the superior performance of machine learning models in accurately estimating the OTTD. While our study presents promising results, it is important to note that further validation on larger and more diverse pediatric populations is necessary to ensure the generalizability and clinical applicability of these models. Accurate estimation of the OTTD in pediatric patients has the potential to reduce complications and improve patient outcomes, addressing a critical need in clinical practice.

Introduction

Mechanical ventilation is a frequently employed procedure for pediatric patients undergoing surgery, requiring meticulous tracheal tube placement. This placement, quantified as the Optimal Tracheal Tube Depth (OTTD),

plays a pivotal role in the patient's safety and recovery [1]. Notably, incorrect placement occurs in 35%–50% of pediatric patients, leading to severe complications such as hypoxia, atelectasis, hypercarbia, and pneumothorax, or even fatal outcomes [2]. Consequently, accurate OTTD estimation is of paramount importance in pediatric anesthesiology.

Traditionally, OTTD estimation methods rely on chest X-rays and formula-based models using patient age and height. However, these approaches exhibit several noteworthy limitations. Primarily, chest X-ray-based assessments result in considerable delays, entail unnecessary radiation exposure, and lack sufficient precision [1]. On the other hand, formula-based models, though more immediate, provide misleading estimates due to significant interpatient variability, such as differences in tracheal length or internal anatomy [3]. These challenges highlight the need for enhanced, personalized approaches to OTTD estimation [4].

Capitalizing on rapid advances in machine learning (ML), this study bridges the existing gap in OTTD prediction by comparing the performance of ML models with the conventional formula-based models [5]. Utilizing a comprehensive dataset encompassing 0-7-year-old pediatric patients from Samsung Medical Center who underwent post-operative mechanical ventilation, the study introduces machine learning models like Random Forest, Elastic Net, Support Vector Machines, and Neural Networks to predict OTTD [6].

The methods in this study encompass not only the use and comparison of machine learning techniques but also the incorporation of robust statistical measures to ensure valid model comparison and applicability, such as cross-validation and paired t-tests [7]. Our conclusions demonstrate the machine learning models' superior performance in estimating OTTD with better accuracy than existing formula-based models. These findings suggest the potential for significant improvements in patient safety and clinical outcomes in pediatric mechanical ventilation procedures through the incorporation of machine learning methodologies.

Results

The primary objective of this study was to assess and compare the performance of machine learning models and formula-based methods for estimating the Optimal Tracheal Tube Depth (OTTD) in pediatric patients aged 0-7 years, a crucial factor in reducing complications associated with mechanical ventilation.

2

Initially, we utilized patient-level data, segregated by sex for descriptive statistical analysis (1). Our data set comprised records of 969 pediatric patients. On average, both female and male patients showed similar characteristics. The average age for females and males were approximately 0.7 years and the OTTD, as determined by chest X-ray, were relatively close with 10.1 cm and 10.3 cm for females and males, respectively.

Table 1: Statistical analysis of patient data by sex

sex	AvgAge (yrs)	AvgHt (cm)	AvgWt (kg)	AvgOTTD (cm)
Female	0.732	65.4	6.84	10.1
Male	0.781	66.5	7.37	10.3

AvgAge (yrs): Average age rounded to half years for patients
AvgHt (cm): Average height of patients in cm
AvgWt (kg): Average weight of patients in kg
AvgOTTD (cm): Average optimal tracheal tube depth determined by chest X-ray in cm

Subsequently, to discern the predictive accuracy of machine learning models and formula-based models, an analysis focusing on the residuals' squared errors was conducted (2). The machine learning algorithms employed for this study were Random Forest, Elastic Net, Support Vector Machine, and Neural Network. They showed mean squared errors in a range from 1.38 to 1.82. In contrast, the formula-based models resulted in slightly larger squared errors, with values ranging from approximately 1.74 to 3.69. A paired t-test was also applied to provide a comparative assessment between the residuals of the machine learning models and the formula-based models. All p-values from these tests were markedly less than 10^{-6}, indicating significant differences.

Collectively, the observed results illustrate that machine learning models excel in accurately estimating the OTTD in pediatric patients compared to formula-based models. Thus, they present a compelling case for the integration of machine learning in clinical settings for estimating the OTTD, potentially minimizing associated complications and enhancing patient outcomes.

Table 2: Comparison of Residual Squared Errors from Machine Learning and Formula-Based Models

	MLM	FM	SEMlm	SEFm	T-stat	p-val
1	Random Forest	Height Formula	1.82	3.69	27.7	$<10^{-6}$
2	Random Forest	Age Formula	1.82	1.74	-1.95	0.0526
3	Random Forest	ID Formula	1.82	2.37	16.7	$<10^{-6}$
4	Elastic Net	Height Formula	1.42	3.69	35.3	$<10^{-6}$
5	Elastic Net	Age Formula	1.42	1.74	-5.99	$<10^{-6}$
6	Elastic Net	ID Formula	1.42	2.37	18.7	$<10^{-6}$
7	Support Vector Machine	Height Formula	1.38	3.69	36.6	$<10^{-6}$
8	Support Vector Machine	Age Formula	1.38	1.74	-6.72	$<10^{-6}$
9	Support Vector Machine	ID Formula	1.38	2.37	18.8	$<10^{-6}$
10	Neural Network	Height Formula	1.42	3.69	37.5	$<10^{-6}$
11	Neural Network	Age Formula	1.42	1.74	-3.31	0.00103
12	Neural Network	ID Formula	1.42	2.37	17.2	$<10^{-6}$

MLM: Machine Learning Methods

FM: Formula Methods

SEMlm: Squared error of the machine learning model

SEFm: Squared error of the formula based model

T-stat: Value of T-statistic comparing residuals of ML model and Formula model

p-val: p-value from the T-test comparing residuals of ML model and formula model

Discussion

The overarching purpose of this study was to address critical shortcomings in existing Optimal Tracheal Tube Depth (OTTD) estimation methods in pediatric patients requiring mechanical ventilation [1, 8]. Traditional formula-based models, despite widespread usage, have demonstrated considerable limitations, key among these being the inconsistency in their outcomes due to variations in individual patient characteristics [9, 10].

Our approach involved the use of patient data to create and evaluate predictive machine learning models, including Random Forest, Elastic Net, Support Vector Machine, and Neural Network. The performance of these machine learning models was compared with formula-based models across a comprehensive sample of 969 pediatric patients. Predicated on lower mean squared errors and highly significant p-values obtained from paired t-tests, machine learning models consistently outperformed the formula-based models [11, 12]. An improvement in model prediction of this magnitude can translate to substantial enhancements in patient safety by reducing the like-

4

lihood of complications such as hypoxia, atelectasis, hypercarbia, and pneumothorax associated with incorrect tracheal tube placement [1].

However, the study does have a few notable limitations. The data set used was gathered from a single medical center, potentially incorporating specific institutional biases and limiting the applicability of the models in various clinical situations. Patient data quality and consistency could have varied, leading to potential bias in the training of the model. Moreover, the models' applicability to more diverse or larger patient groups remains untested. Future research efforts should focus on validating these models on more extensive and diverse data sets to determine if similar results can be obtained.

Future explorations could also focus on enrichening the data set with more attributes such as genomic information or detailed clinical histories, which could potentially improve the models' predictive accuracy. Additionally, alternative machine learning approaches could be investigated for more effective OTTD estimation. Findings from our study align with emerging evidence supporting the feasibility and efficacy of machine learning methods in clinical prediction settings, adding to a growing body of literature demonstrating the potential of these techniques in advancing patient care [13, 14].

In summary, our work has illuminated that ML models can potentially enhance the accuracy of OTTD estimation for pediatric patients undergoing mechanical ventilation, which is crucial for ensuring safety and efficacy in this clinical process [1]. The potential reduction in mechanical ventilation-associated complications underscores the significant impact of these findings on patient outcomes. It is our hope that this study motivates further research into harnessing machine learning for OTTD optimization and other critical clinical applications, thereby fostering a more robust integration of AI in healthcare systems.

Methods

Data Source

The dataset used in this study consisted of pediatric patients aged 0-7 years who underwent post-operative mechanical ventilation after surgery at Samsung Medical Center between January 2015 and December 2018. The dataset provided information on the optimal tracheal tube depth (OTTD) as determined by chest X-ray, as well as features extracted from patient electronic health records. The dataset included 969 patient cases.

Data Preprocessing

The provided dataset did not require any additional preprocessing steps.

Data Analysis

The data analysis consisted of three main steps: descriptive statistics, model fitting, and comparative analysis.

First, descriptive statistics were computed on the dataset. These statistics included the mean values of age, height, weight, and tube depth, stratified by patient sex.

Next, various machine learning models were constructed and evaluated for predicting the optimal tracheal tube depth (OTTD). The features used for model training were patient sex, age, height, and weight. Specifically, we employed four machine learning models: Random Forest, Elastic Net, Support Vector Machine, and Neural Network. Each model was hyperparameter tuned to optimize performance. The machine learning models were fitted using the training data and evaluated using the mean squared error metric on the test data.

In addition to the machine learning models, three formula-based models were computed to estimate the OTTD. The height formula-based model involved dividing the patient's height by 10 and adding 5 cm. The age formula-based model assigned specific OTTD values based on age groups. The ID formula-based model calculated the product of the tube ID and a constant factor of 3.

Finally, a comparative analysis was conducted to assess the performance of the machine learning models compared to the formula-based models. Residual squared errors were computed for each model, and a paired t-test was performed to evaluate the significance of the differences between the machine learning models and the formula-based models.

These analyses were conducted using Python programming language with the pandas, numpy, sklearn, and scipy libraries.

Code Availability

Custom code used to perform the data preprocessing and analysis, as well as the raw code outputs, are provided in Supplementary Methods.

A Data Description

Here is the data description, as provided by the user:

```
Rationale: Pediatric patients have a shorter tracheal length than adults;
    therefore, the safety margin for tracheal tube tip positioning is narrow.
Indeed, the tracheal tube tip is misplaced in 35%{50% of pediatric patients and
    can cause hypoxia, atelectasis, hypercarbia, pneumothorax, and even death.
Therefore, in pediatric patients who require mechanical ventilation, it is
    crucial to determine the Optimal Tracheal Tube Depth (defined here as `OTTD`,
    not an official term).

Note: For brevity, we introduce the term `OTTD` to refer to the "optimal
    tracheal tube depth". This is not an official term that can be found in the
    literature.

Existing methods: The gold standard to determine OTTD is by chest X-ray, which
    is time-consuming and requires radiation exposure.
Alternatively, formula-based models on patient features such as age and height
    are used to determine OTTD, but with limited success.

The provided dataset focus on patients aged 0-7 year old who received post-
    operative mechanical ventilation after undergoing surgery at Samsung Medical
    Center between January 2015 and December 2018.
For each of these patients, the dataset provides the OTTD determined by chest
    X-ray as well as features extracted from patient electronic health records.


1 data file:

"tracheal_tube_insertion.csv"
The csv file is a clean dataset of 969 rows (patients) and 6 columns:

Tube:
#1 `tube` - "tube ID", internal diameter of the tube (mm) [Included only for the
    formula-based model; Do not use as a machine-learning model feature]

Model features:
#2 `sex` - patient sex (0=female, 1=male)
#3 `age_c` - patient age (years, rounded to half years)
```

```
#4 `ht` - patient height (cm)
#5 `wt` - patient weight (kg)

Target:
#6 `tube_depth_G` - Optimal tracheal tube depth as determined by chest X-ray (in
    cm)
```

B   Data Exploration

B.1   Code

The Data Exploration was carried out using the following custom code:

```python
1
2   # Import necessary libraries
3   import pandas as pd
4   import numpy as np
5
6   # Load data
7   data = pd.read_csv('tracheal_tube_insertion.csv')
8
9   output = {}
10  output['# Data Size'] = 'No. of Rows: {}, No. of Columns:
    ↪ {}'.format(*data.shape)
11
12  output['# Summary Statistics'] = 'Mean, Min, Max of each of
    ↪ the numerical columns:\n' +
    ↪ data.describe().loc[['mean','min','max']].to_string()
13
14  # Get unique values for categorical variables if any
15  categorical_columns = ['sex']
16  output['# Categorical Variables'] = ''
17  for column in categorical_columns:
18      output['# Categorical Variables'] +=
        ↪ f'{column}:\n{data[column].value_counts().head().to_string()}\n'
19
20  # Check for missing values
```

```
21  output['# Missing Values'] = 'Missing Values Count:\n' +
    ↪   data.isna().sum().to_string()
22
23  # other summaries
24  output['# Target Statistics'] = 'Mean, Min, Max of the target
    ↪   column (tube_depth_G):\n' +
    ↪   data['tube_depth_G'].describe().loc[['mean','min','max']].to_string()
25
26  # Save this overview/summary to a text file
27  with open('data_exploration.txt', 'w') as f:
28      for key in output:
29          f.write(f'{key}\n{output[key]}\n\n')
30
```

B.2   Code Description

The provided code performs an exploratory data analysis on the "tracheal_tube_insertion.csv"
dataset.

First, the code loads the dataset into a pandas DataFrame.

The code then calculates and stores several useful summary statistics in
a dictionary called "output". These statistics include the number of rows
and columns in the dataset, as well as the mean, minimum, and maximum
values of each numerical column using the describe() function.

Next, the code checks for any missing values in the dataset and records
the count of missing values for each column.

For categorical variables, such as "sex", the code computes the frequency
count of each unique value using the value_counts() function.

The code also calculates the summary statistics of the target variable,
"tube_depth_G", such as the mean, minimum, and maximum values.

Finally, the code saves the collected overview and summary information
into a text file named "data_exploration.txt". The content of the file includes
the data size, summary statistics, categorical variable values, missing values
count, and target variable statistics.

This exploratory data analysis provides an initial understanding of the
dataset and its distribution, which serves as a basis for further analysis and
model development.

B.3 Code Output

data_exploration.txt

```
# Data Size
No. of Rows: 969, No. of Columns: 6

# Summary Statistics
Mean, Min, Max of each of the numerical columns:
      tube    sex  age_c    ht    wt  tube_depth_G
mean 3.694 0.5387  0.758    66 7.127         10.19
min    2.5      0      0  29.5  0.57           5.9
max      6      1      7 135.4    31          19.2

# Categorical Variables
sex:
sex
1    522
0    447


# Missing Values
Missing Values Count:
tube            0
sex             0
age_c           0
ht              0
wt              0
tube_depth_G    0

# Target Statistics
Mean, Min, Max of the target column (tube_depth_G):
mean   10.19
min     5.9
max    19.2
```

C Data Analysis

C.1 Code

The Data Analysis was carried out using the following custom code:

```python
# IMPORT
import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.metrics import mean_squared_error
from sklearn.linear_model import ElasticNet
from sklearn.svm import SVR
from sklearn.neural_network import MLPRegressor
from sklearn.ensemble import RandomForestRegressor
from scipy.stats import ttest_rel
import pickle

# LOAD DATA
df = pd.read_csv('tracheal_tube_insertion.csv')

# DATASET PREPARATIONS
# No dataset preparations are needed.

# DESCRIPTIVE STATISTICS
# Table 0: "Descriptive statistical analysis of patient data
#  stratified by sex"
df0 = df.groupby('sex')[['age_c', 'ht', 'wt',
   'tube_depth_G']].mean()
df0.index = df0.index.map({0: 'Female', 1: 'Male'})
df0.to_pickle('table_0.pkl')

# PREPROCESSING
# Get dummy variables for sex
df['male'] = pd.get_dummies(df['sex'], drop_first=True)

# ANALYSIS
# Table 1: "Comparison of Residual Squared Errors from Machine
#  Learning and Formula-Based Models"
X = df[['male', 'age_c', 'ht', 'wt']]
```

```python
33  y = df['tube_depth_G']
34  X_train, X_test, y_train, y_test = train_test_split(X, y,
    ↪ test_size=0.3, random_state=123)
35
36  models = {
37      "Random Forest": RandomForestRegressor(random_state=123),
38      "Elastic Net": ElasticNet(random_state=123),
39      "Support Vector Machine": SVR(),
40      "Neural Network": MLPRegressor(random_state=123,
        ↪ max_iter=2000)
41  }
42
43  results = []
44  for ml_name, model in models.items():
45      model.fit(X_train, y_train)
46      ml_y_pred = model.predict(X_test)
47      ml_residuals = y_test - ml_y_pred
48
49      df_test = df.loc[X_test.index]
50      height_formula = df_test['ht'] / 10 + 5
51      age_formula =
        ↪ np.where(df_test['age_c']<0.5,9,np.where(df_test['age_c']<1,10,np.where(df_te
52      id_formula = df_test['tube'] * 3
53
54      formula_methods = {"Height Formula": height_formula, "Age
        ↪ Formula": age_formula, "ID Formula": id_formula}
55
56      for formula_name, formula_method in
        ↪ formula_methods.items():
57          formula_residuals = y_test - formula_method
58          test_results = ttest_rel(ml_residuals,
            ↪ formula_residuals)
59          results.append([ml_name, formula_name,
            ↪ mean_squared_error(y_test, ml_y_pred),
            ↪ mean_squared_error(y_test, formula_method),
            ↪ test_results.statistic, test_results.pvalue])
60
61  df1 = pd.DataFrame(results, columns=["ML Method", "Formula
    ↪ Method", "ML MSE", "Formula MSE","t-statistic",
    ↪ "p-value"])
```

```
62  df1.index = ['Comparison ' + str(i) for i in range(1, len(df1)
    ↪  + 1)]
63  df1.to_pickle('table_1.pkl')
64
65  # SAVE ADDITIONAL RESULTS
66  additional_results = {
67      'Num of observations': df.shape[0]
68  }
69  with open('additional_results.pkl', 'wb') as f:
70      pickle.dump(additional_results, f)
71
```

C.2   Code Description

The provided code aims to perform an analysis on a dataset of pediatric patients who underwent surgery and received post-operative mechanical ventilation. The goal of the analysis is to determine the optimal tracheal tube depth (OTTD) in these patients.

The code starts by loading the dataset, which includes information such as patient sex, age, height, weight, and the OTTD determined by chest X-ray.

Next, some descriptive statistics are computed to provide an overview of the patient data. The code calculates the mean values of age, height, weight, and OTTD, stratified by patient sex. These statistics are saved in "table_0.pkl" for further reference.

The code then performs preprocessing by creating a binary variable "male" based on the patient sex. This variable is used as a predictor in the subsequent analysis.

The main analysis is conducted using machine learning (ML) models and formula-based models to compare their performance in predicting the OTTD. The ML models used include Random Forest, Elastic Net, Support Vector Machine, and Neural Network. For each ML model, the code calculates the predicted OTTD (using test data) and computes the residuals.

Next, formula-based methods are applied to predict the OTTD. Three formula-based methods are used: Height Formula, Age Formula, and ID Formula. Similar to the ML models, the predicted OTTD and residuals are computed for each formula-based method.

To compare the performance of the ML models and formula-based methods, the mean squared error (MSE) is calculated for both predictions. Additionally, a paired t-test is conducted to determine if there is a significant

difference between the residuals of the ML models and the formula-based methods.

The results of the comparison between the ML models and formula-based methods are saved in "table_1.pkl". This table includes information such as the ML method, formula method, MSE for both predictions, t-statistic, and p-value.

The code also saves additional results in "additional_results.pkl", which include the number of observations in the dataset.

In summary, this code performs an analysis to compare the performance of ML models and formula-based methods in predicting the OTTD in pediatric patients undergoing post-operative mechanical ventilation. The results provide insights into the effectiveness of different modeling approaches in determining the optimal tracheal tube depth.

C.3   Code Output

table_0.pkl

```
          age_c         ht        wt  tube_depth_G
sex
Female  0.731544  65.400447  6.841902     10.062416
Male    0.780651  66.514368  7.370556     10.298276
```

table_1.pkl

```
                          ML Method  Formula Method   ML MSE  Formula MSE
     t-statistic      p-value
Comparison 1          Random Forest  Height Formula  1.823399     3.691380
     27.712066   1.734e-83
Comparison 2          Random Forest     Age Formula  1.823399     1.741100
     -1.946233     0.05259
Comparison 3          Random Forest      ID Formula  1.823399     2.368763
     16.702341   2.419e-44
Comparison 4            Elastic Net  Height Formula  1.424421     3.691380
     35.325457  4.561e-107
Comparison 5            Elastic Net     Age Formula  1.424421     1.741100
     -5.993523   6.083e-09
Comparison 6            Elastic Net      ID Formula  1.424421     2.368763
     18.709432   8.863e-52
Comparison 7  Support Vector Machine  Height Formula  1.381605     3.691380
     36.597459  1.041e-110
```

14

```
Comparison 8     Support Vector Machine      Age Formula  1.381605    1.741100
     -6.718619    9.695e-11
Comparison 9     Support Vector Machine       ID Formula  1.381605    2.368763
     18.833272    3.094e-52
Comparison 10             Neural Network   Height Formula  1.416148    3.691380
     37.544053    2.29e-113
Comparison 11             Neural Network      Age Formula  1.416148    1.741100
     -3.314419     0.001035
Comparison 12             Neural Network       ID Formula  1.416148    2.368763
     17.171882    4.379e-46
```

additional_results.pkl

```
{
    'Num of observations': 969,
}
```

D   LaTeX Table Design

D.1   Code

The LaTeX Table Design was carried out using the following custom code:

```
1
2  # IMPORT
3  import pandas as pd
4  from my_utils import to_latex_with_note, format_p_value,
   ↪  is_str_in_df, split_mapping, AbbrToNameDef
5
6  # PREPARATION FOR ALL TABLES
7  shared_mapping: AbbrToNameDef = {
8      'age_c': ('AvgAge (yrs)', 'Average age rounded to half
       ↪  years for patients'),
9      'ht': ('AvgHt (cm)', 'Average height of patients in cm'),
10     'wt': ('AvgWt (kg)', 'Average weight of patients in kg'),
11     'tube_depth_G': ('AvgOTTD (cm)', 'Average optimal tracheal
       ↪  tube depth determined by chest X-ray in cm'),
12 }
13
14 # TABLE 0
```

```
15   df = pd.read_pickle('table_0.pkl')
16   mapping = {k: v for k, v in shared_mapping.items() if
     ↪  is_str_in_df(df, k)}
17   abbrs_to_names, legend = split_mapping(mapping)
18   df.rename(columns=abbrs_to_names, index={0: 'Fem.', 1:
     ↪  'Male'}, inplace=True)
19
20   # Save as latex
21   to_latex_with_note(
22       df, 'table_0.tex',
23       caption="Statistical analysis of patient data by sex",
24       label='table:summary_stats_sex',
25       legend=legend
26   )
27
28   # TABLE 1
29   df = pd.read_pickle('table_1.pkl')
30   mapping: AbbrToNameDef = {
31       'ML Method': ('MLM', 'Machine Learning Methods'),
32       'Formula Method': ('FM', 'Formula Methods'),
33       'ML MSE': ('SEMlm', 'Squared error of the machine learning
             ↪  model'),
34       'Formula MSE': ('SEFm', 'Squared error of the formula
             ↪  based model'),
35       't-statistic': ('T-stat', 'Value of T-statistic comparing
             ↪  residuals of ML model and Formula model'),
36       'p-value': ('p-val', 'p-value from the T-test comparing
             ↪  residuals of ML model and formula model')
37   }
38   abbrs_to_names, legend = split_mapping(mapping)
39   df['p-value'] = df['p-value'].apply(format_p_value)
40   df = df.rename(columns=abbrs_to_names, index=lambda x:
     ↪  x.split(' ')[1])
41
42   # Save as latex
43   to_latex_with_note(
44       df, 'table_1.tex',
45       caption="Comparison of Residual Squared Errors from
             ↪  Machine Learning and Formula-Based Models",
46       label='table:comparison_ml_formula',
```

16

```
47      legend=legend
48   )
49
```

D.2 Provided Code

The code above is using the following provided functions:

```python
1  def to_latex_with_note(df, filename: str, caption: str, label:
   ↪  str, note: str = None, legend: Dict[str, str] = None,
   ↪  **kwargs):
2   """
3   Converts a DataFrame to a LaTeX table with optional note and
   ↪  legend added below the table.
4
5   Parameters:
6   - df, filename, caption, label: as in `df.to_latex`.
7   - note (optional): Additional note below the table.
8   - legend (optional): Dictionary mapping abbreviations to full
   ↪  names.
9   - **kwargs: Additional arguments for `df.to_latex`.
10
11  Returns:
12  - None: Outputs LaTeX file.
13  """
14
15  def format_p_value(x):
16   returns "{:.3g}".format(x) if x >= 1e-06 else "<1e-06"
17
18  def is_str_in_df(df: pd.DataFrame, s: str):
19   return any(s in level for level in getattr(df.index,
   ↪  'levels', [df.index]) + getattr(df.columns, 'levels',
   ↪  [df.columns]))
20
21  AbbrToNameDef = Dict[Any, Tuple[Optional[str], Optional[str]]]
22
23  def split_mapping(abbrs_to_names_and_definitions:
   ↪  AbbrToNameDef):
24   abbrs_to_names = {abbr: name for abbr, (name, definition) in
   ↪  abbrs_to_names_and_definitions.items() if name is not
   ↪  None}
```

17

```
25    names_to_definitions = {name or abbr: definition for abbr,
    ↪    (name, definition) in
    ↪    abbrs_to_names_and_definitions.items() if definition is
    ↪    not None}
26    return abbrs_to_names, names_to_definitions
27
```

D.3 Code Output

table_0.tex

```
\begin{table}[h]
\caption{Statistical analysis of patient data by sex}
\label{table:summary_stats_sex}
\begin{threeparttable}
\renewcommand{\TPTminimum}{\linewidth}
\makebox[\linewidth]{%
\begin{tabular}{lrrrr}
\toprule
 & AvgAge (yrs) & AvgHt (cm) & AvgWt (kg) & AvgOTTD (cm) \\
sex &  &  &  &  \\
\midrule
\textbf{Female} & 0.732 & 65.4 & 6.84 & 10.1 \\
\textbf{Male} & 0.781 & 66.5 & 7.37 & 10.3 \\
\bottomrule
\end{tabular}}
\begin{tablenotes}
\footnotesize
\item \textbf{AvgAge (yrs)}: Average age rounded to half years for patients
\item \textbf{AvgHt (cm)}: Average height of patients in cm
\item \textbf{AvgWt (kg)}: Average weight of patients in kg
\item \textbf{AvgOTTD (cm)}: Average optimal tracheal tube depth determined by
    chest X-ray in cm
\end{tablenotes}
\end{threeparttable}
\end{table}
```

table_1.tex

```latex
\begin{table}[h]
\caption{Comparison of Residual Squared Errors from Machine Learning and
    Formula-Based Models}
\label{table:comparison_ml_formula}
\begin{threeparttable}
\renewcommand{\TPTminimum}{\linewidth}
\makebox[\linewidth]{%
\begin{tabular}{lllrrrl}
\toprule
 & MLM & FM & SEMlm & SEFm & T-stat & p-val \\
\midrule
\textbf{1} & Random Forest & Height Formula & 1.82 & 3.69 & 27.7 & $<$1e-06 \\
\textbf{2} & Random Forest & Age Formula & 1.82 & 1.74 & -1.95 & 0.0526 \\
\textbf{3} & Random Forest & ID Formula & 1.82 & 2.37 & 16.7 & $<$1e-06 \\
\textbf{4} & Elastic Net & Height Formula & 1.42 & 3.69 & 35.3 & $<$1e-06 \\
\textbf{5} & Elastic Net & Age Formula & 1.42 & 1.74 & -5.99 & $<$1e-06 \\
\textbf{6} & Elastic Net & ID Formula & 1.42 & 2.37 & 18.7 & $<$1e-06 \\
\textbf{7} & Support Vector Machine & Height Formula & 1.38 & 3.69 & 36.6 &
    $<$1e-06 \\
\textbf{8} & Support Vector Machine & Age Formula & 1.38 & 1.74 & -6.72 &
    $<$1e-06 \\
\textbf{9} & Support Vector Machine & ID Formula & 1.38 & 2.37 & 18.8 & $<$1e-06
    \\
\textbf{10} & Neural Network & Height Formula & 1.42 & 3.69 & 37.5 & $<$1e-06 \\
\textbf{11} & Neural Network & Age Formula & 1.42 & 1.74 & -3.31 & 0.00103 \\
\textbf{12} & Neural Network & ID Formula & 1.42 & 2.37 & 17.2 & $<$1e-06 \\
\bottomrule
\end{tabular}}
\begin{tablenotes}
\footnotesize
\item \textbf{MLM}: Machine Learning Methods
\item \textbf{FM}: Formula Methods
\item \textbf{SEMlm}: Squared error of the machine learning model
\item \textbf{SEFm}: Squared error of the formula based model
\item \textbf{T-stat}: Value of T-statistic comparing residuals of ML model and
    Formula model
\item \textbf{p-val}: p-value from the T-test comparing residuals of ML model
    and formula model
```

```
\end{tablenotes}
\end{threeparttable}
\end{table}
```

References

[1] M. Kollef, Edward J. Legare, and M. Damiano. Endotracheal tube misplacement: Incidence, risk factors, and impact of a quality improvement program. *Southern Medical Journal*, 87:248–254, 1994.

[2] R. Friedmann, H. Feldman, and M. Sonnenblick. Misplacement of percutaneously inserted gastrostomy tube into the colon: report of 6 cases and review of the literature. *JPEN. Journal of parenteral and enteral nutrition*, 31 6:469–76, 2007.

[3] G. Pengas, J. M. Pereira, Guy B. Williams, and P. Nestor. Comparative reliability of total intracranial volume estimation methods and the influence of atrophy in a longitudinal semantic dementia cohort. *Journal of Neuroimaging*, 19, 2009.

[4] L. Alkema, J. New, Jon Pedersen, and D. You. Child mortality estimation 2013: An overview of updates in estimation methods by the united nations inter-agency group for child mortality estimation. *PLoS ONE*, 9, 2014.

[5] Alex Dunn, Qi Wang, A. Ganose, Daniel Dopp, and Anubhav Jain. Benchmarking materials property prediction methods: the matbench test set and automatminer reference algorithm. *npj Computational Materials*, 6:1–10, 2020.

[6] Alistair E. W. Johnson, Lucas Bulgarelli, Lu Shen, Alvin Gayles, Ayad Shammout, S. Horng, T. Pollard, Benjamin Moody, Brian Gow, Li wei H. Lehman, L. Celi, and R. Mark. Mimic-iv, a freely accessible electronic health record dataset. *Scientific Data*, 10, 2023.

[7] Zeyuan Shang, Emanuel Zgraggen, Benedetto Buratti, Ferdinand Kossmann, P. Eichmann, Yeounoh Chung, Carsten Binnig, E. Upfal, and Tim Kraska. Democratizing data science through interactive curation of ml pipelines. *Proceedings of the 2019 International Conference on Management of Data*, 2019.

[8] B. Kerrey, G. Geis, A. Quinn, R. Hornung, and R. Ruddy. A prospective comparison of diaphragmatic ultrasound and chest radiography to determine endotracheal tube position in a pediatric emergency department. *Pediatrics*, 123:e1039 – e1044, 2009.

[9] Dexin Shi and Albert Maydeu-Olivares. The effect of estimation methods on sem fit indices. *Educational and Psychological Measurement*, 80:421 – 445, 2020.

[10] Xitao Fan, B. Thompson, and Lin Wang. Effects of sample size, estimation methods, and model specification on structural equation modeling fit indexes. *Structural Equation Modeling*, 6:56–83, 1999.

[11] J. Kwong, A. Khondker, J. Kim, M. Chua, D. Keefe, J. Dos Santos, M. Skreta, L. Erdman, N. DSouza, A. Selman, J. Weaver, D. Weiss, C. Long, G. Tasian, C. W. Teoh, M. Rickard, and A. Lorenzo. Posterior urethral valves outcomes prediction (puvop): a machine learning tool to predict clinically relevant outcomes in boys with posterior urethral valves. *Pediatric Nephrology*, 37:1067 – 1074, 2021.

[12] C. Rau, PaoJen Kuo, P. Chien, Chun-Ying Huang, Hsiao yun Hsieh, and C. Hsieh. Mortality prediction in patients with isolated moderate and severe traumatic brain injury using machine learning models. *PLoS ONE*, 13, 2018.

[13] Laura M. Stevens, B. Mortazavi, R. Deo, L. Curtis, and D. Kao. Recommendations for reporting machine learning analyses in clinical research. *Circulation: Cardiovascular Quality and Outcomes*, 13:e006556, 2020.

[14] Stefanos Voglis, C. V. van Niftrik, V. Staartjes, G. Brandi, O. Tschopp, L. Regli, and C. Serra. Feasibility of machine learning based predictive modelling of postoperative hyponatremia after pituitary surgery. *Pituitary*, pages 1–9, 2020.