## You

Traceback (most recent call last):
File "C:\Users\fifau\OneDrive - University of Bradford\Documents\Python\_Files\Funzy ML\DrinksSalesML.py", line 18, in
model.fit(X\_train, y\_train)
File "C:\Users\fifau\AppData\Local\Programs\Python\Python311\Lib\site-packages\sklearn\base.py", line 1152, in wrapper
return fit\_method(estimator, \*args, \*\*kwargs)
^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
File "C:\Users\fifau\AppData\Local\Programs\Python\Python311\Lib\site-packages\sklearn\linear\_model\_base.py", line 678, in fit
X, y = self.\_validate\_data(
^^^^^^^^^^^^^^^^^^^^
File "C:\Users\fifau\AppData\Local\Programs\Python\Python311\Lib\site-packages\sklearn\base.py", line 622, in \_validate\_data
X, y = check\_X\_y(X, y, \*\*check\_params)
^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
File "C:\Users\fifau\AppData\Local\Programs\Python\Python311\Lib\site-packages\sklearn\utils\validation.py", line 1146, in check\_X\_y
X = check\_array(
^^^^^^^^^^^^
File "C:\Users\fifau\AppData\Local\Programs\Python\Python311\Lib\site-packages\sklearn\utils\validation.py", line 915, in check\_array
array = \_asarray\_with\_order(array, order=order, dtype=dtype, xp=xp)
^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
File "C:\Users\fifau\AppData\Local\Programs\Python\Python311\Lib\site-packages\sklearn\utils\_array\_api.py", line 380, in \_asarray\_with\_order
array = numpy.asarray(array, order=order, dtype=dtype)
^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
File "C:\Users\fifau\AppData\Local\Programs\Python\Python311\Lib\site-packages\pandas\core\generic.py", line 2084, in **array**
arr = np.asarray(values, dtype=dtype)
^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
ValueError: could not convert string to float: 'Tue 25/6/24 10:00-11:00'

## Copilot

It looks like your model is trying to convert a string to a float, which isn't possible directly. The string 'Tue 25/6/24 10:00-11:00' seems to be a date-time value. You'll need to preprocess this data into a numerical format before fitting your model.

One common approach is to convert date-time strings into timestamps or extract relevant features like hour of the day, day of the week, etc., as separate numerical columns. Would you like guidance on how to do this?

## You

Yes how can I do it so that the relevant information that is preserved is the day and time interval?

## Copilot

You can use the pandas library to extract the day of the week and time interval from your date-time string. Here's a brief example of how you might do it:

import pandas as pd

# Assuming df is your DataFrame and 'date\_column' is the name of your date-time column

df['date\_column'] = pd.to\_datetime(df['date\_column'], format='%a %d/%m/%y %H:%M-%H:%M')

# Extract day of week as a numerical value (Monday=0, Sunday=6)

df['day\_of\_week'] = df['date\_column'].dt.dayofweek

# Extract start hour and end hour from the time interval

df['start\_hour'] = df['date\_column'].dt.hour

df['end\_hour'] = df['date\_column'].dt.hour + 1 # Assuming the interval is always 1 hour

This code converts your date-time string into a pandas datetime object, from which you can extract the day of the week and start/end hours as separate columns. You'll need to adjust the format parameter in pd.to\_datetime() to match the exact format of your date-time strings.

Remember to drop or convert any other non-numeric columns before fitting your model. Would you like more detailed help with this?

## You

Now I'm getting this error: C:\Users\fifau\AppData\Local\Programs\Python\Python311\python.exe "C:\Users\fifau\OneDrive - University of Bradford\Documents\Python\_Files\Funzy ML\DrinksSalesML.py"
Traceback (most recent call last):
File "C:\Users\fifau\OneDrive - University of Bradford\Documents\Python\_Files\Funzy ML\DrinksSalesML.py", line 14, in
data['Date/Time'] = pd.to\_datetime(data['Date/Time'], format='%a %d/%m/%y %H:%M-%H:%M')
^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
File "C:\Users\fifau\AppData\Local\Programs\Python\Python311\Lib\site-packages\pandas\core\tools\datetimes.py", line 1112, in to\_datetime
values = convert\_listlike(arg.\_values, format)
^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
File "C:\Users\fifau\AppData\Local\Programs\Python\Python311\Lib\site-packages\pandas\core\tools\datetimes.py", line 488, in \_convert\_listlike\_datetimes
return \_array\_strptime\_with\_fallback(arg, name, utc, format, exact, errors)
^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
File "C:\Users\fifau\AppData\Local\Programs\Python\Python311\Lib\site-packages\pandas\core\tools\datetimes.py", line 519, in \_array\_strptime\_with\_fallback
result, timezones = array\_strptime(arg, fmt, exact=exact, errors=errors, utc=utc)
^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
File "strptime.pyx", line 238, in pandas.\_libs.tslibs.strptime.array\_strptime
File "strptime.pyx", line 248, in pandas.*libs.tslibs.strptime.array\_strptime
File "C:\Users\fifau\AppData\Local\Programs\Python\Python311\Lib\_strptime.py", line 263, in compile
return re\_compile(self.pattern(format), IGNORECASE)
^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
File "C:\Users\fifau\AppData\Local\Programs\Python\Python311\Lib\re\_init*.py", line 227, in compile
return *compile(pattern, flags)
^^^^^^^^^^^^^^^^^^^^^^^^
File "C:\Users\fifau\AppData\Local\Programs\Python\Python311\Lib\re\_init*.py", line 294, in \_compile
p = \_compiler.compile(pattern, flags)
^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
File "C:\Users\fifau\AppData\Local\Programs\Python\Python311\Lib\re\_compiler.py", line 743, in compile
p = \_parser.parse(p, flags)
^^^^^^^^^^^^^^^^^^^^^^^
File "C:\Users\fifau\AppData\Local\Programs\Python\Python311\Lib\re\_parser.py", line 980, in parse
p = \_parse\_sub(source, state, flags & SRE\_FLAG\_VERBOSE, 0)
^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
File "C:\Users\fifau\AppData\Local\Programs\Python\Python311\Lib\re\_parser.py", line 455, in \_parse\_sub
itemsappend(\_parse(source, state, verbose, nested + 1,
^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
File "C:\Users\fifau\AppData\Local\Programs\Python\Python311\Lib\re\_parser.py", line 858, in \_parse
raise source.error(err.msg, len(name) + 1) from None
re.error: redefinition of group name 'H' as group 7; was group 5 at position 167

Process finished with exit code 1

## Copilot

The error you're encountering now is related to the format string used in pd.to\_datetime(). It seems there might be an issue with the way the hours are specified in the format.

In Python's datetime module, %H is used to parse hours (00-23). If you have two hour specifications in your date-time string (like '10:00-11:00'), you'll need to handle them separately. However, you cannot have two groups with the same name in a regular expression, which is likely causing the error.

Here's how you can adjust your code:

data['start\_time'] = pd.to\_datetime(data['Date/Time'].str.split('-').str[0], format='%a %d/%m/%y %H:%M')

data['end\_time'] = pd.to\_datetime(data['Date/Time'].str.split('-').str[1], format='%H:%M')

This code splits the 'Date/Time' column on the '-' character and parses the start and end times separately. Make sure your date-time strings are consistent and match the format you're providing to pd.to\_datetime().

## You

/i have this code but I'm getting this error:

## Copilot

Could you please provide