def create\_chatcompletion\_messages(row):

return [

{

"role": "system",

"content": "As a Biological Data Analyst, your primary task is to standardize the 'Treatment\_Duration' field by converting durations and ensuring consistent, clear formatting across all entries."

},

{

"role": "user",

"content": f"Carefully review the info below: {row['Treatment\_Duration']}\n"

"-Before starting the standardization, check the total number of Original Terms provided. Ensure that the order of terms in your Standardized terms matches the exact order in the input data. Standardize each Original Term sequentially without omission, and verify that the total number of Standardized Terms matches the count in the input.\n"

"-Preserve Composite Terms: Keep complex duration terms as they are to maintain detail, e.g., '14 hours + 5 days + 3 hours + 1 minute' remains unchanged. Convert'1 day + 4 day' to '1 Day + 4 Days'. Do not add as '5 Days'. Consider '48h,24h,6h' as one single term.\n"

"-Format ranges with hyphens and no spaces for clarity. Ensure proper capitalization."

"e.g., 'from 4 weeks to 6 months' becomes '4 Weeks-6 Months'.\n"

"'10 to 12 weeks'becomes '10-12 Weeks'\n."

"-Standardize Approximations: Standardize approximate times by removing spaces next to the tilde, e.g., '~ 18 Hours' becomes '~18 Hours'.\n"

"-Singular and Plural Terms for 0 terms: Standardize to singular where applicable, e.g., '0 Hours' becomes '0 Hour'. Also, ensure uniformity in terminology, changing 'LT' to 'Long-Term'.\n"

"-Consistent Formatting: Capitalize the first letter of each term and maintain consistent terms across entries. For instance, 'overnight + 0 Hours' should be 'Overnight + 0 Hour'.\n"

"- Special Terms Standardization: Standardize terms like 'From birth' to 'From Birth' and 'From day 1 to day 28' to 'From Day 1 to Day 28'.\n"

"-Remove the ~ symbol, for example, ~18 Hours convert to 18 Hours"

"Format the Output as a Table:\n"

"Create a table with two columns:\n"

"Column 1: List each Original Term exactly as it appears in the input. Preserve Original Term and do not apply any formatting or transformations to the original input term. Simply copy it directly into the first column of the output.\n"

"Column 2: Provide the standardized term for each corresponding original term.\n\n"

"Apply biological nomenclature standards and domain expertise to ensure accuracy in the duration standardization process."

"Do not include any explanations in any format."

}

]