

Translation of Units dataset – Guidelines

Goal:

- The goal is to translate the words used for measurement present in the dataset, closest to how they are culturally used in the native language.
- If the unit does not have a corresponding word in the language, you may transliterate.

Instructions:

- The dataset contains 901 sentences.
- The dataset contains sentences which are complex, hence the use of Machine Translation and Dictionary is allowed, but the outputs should not be copy-pasted.
- In the numerical.txt section, the dataset contains numbers, fractions in numbers (like $\frac{1}{2}$) or words (like 2 out of 5), alphanumerics (1st, 2nd etc.) or number in text form (twenty, first etc) which must be translated.

Basic examples of translation English to Telugu:

Example1:

Src: This road is 10 **kilometers** long and leads to Agra.

Tgt1: ఈ రోడు 10 **కిలోమీటర్ల** పొడవు ఉండి ఆగ్రాకు దారితేస్తుంది.

Tgt2: यह सड़क 10 किलोमीटर लंबी है और आगरा तक जाती है।

Explanation:

Tgt1: For the word **Kilometer** there is no native word in Telugu, hence the use of the word కిలోమీటర్ (read as kilometer-la meaning 'kilometer's' which is a plural form).

Tgt2: For the word **Kilometer** there is no native word in Hindi, hence the use of the word किलोमीटर (read as kilometer)

Example2:

Src: If I grow one more inch, I will be 6 feet.

Tgt1: నేను ఇంచ్ అంగుళం పెరిగితే, 6 అడుగులు ఉంటాను.

Tgt2: अगर मैं एक इंच और बढ़ जाऊं तो मैं 6 फीट का हो जाऊंगा।

Explanation:

Tgt1: Here both the words **inch** and **feet** have a native term in Telugu, which are

- Inch - అంగుళం (read as "Angulam")
- Feet - అడుగులు (read as "Adugulu"). అడుగు - Adugu meaning foot used for measurement + లు which is plural suffix).

Tgt2: There are no native Hindi words for the terms "inch" and "feet", hence the use of the terms इंच (read as inch) and फीट (read as feet).

Example3:

Src: I was going to work **yesterday**, when I saw Anand.

Tgt1: నేను నిన్న పనికి వెళ్లున్నప్పుడు ఆనంద్ ని చూశాను.

Tgt2: मैं कल काम पर जा रहा था, तभी मैंने आनंद को देखा।

Explanation:

Tgt1: There is an existing term for the word **yesterday**, which is నిన్న (read as Ninna), which is the exact translation.

Tgt2: There is an existing term for the word **yesterday**, which is कल (read as Kal), which is the exact translation.

Example4:

Src: I **almost** got hit by a car.

Tgt1: అ శారు నన్ను దాదాపు దీకొట్టింది.

Tgt2: मैं लगभग एक कार से टकरा गया था।

Explanation:

Tgt1: The word **almost** has a native term in Telugu which is దాదాపు (Daa-daa-pu where “D” sounds like “th” in the word “this”).

Tgt2: The word **almost** has a native term in Hindi which is लगभग (Read as lagh-bhag).

Example5:

Src: After the party, he left our house in the **blink of an eye**.

Tgt1: పార్టీ తర్వాత, అతను రెప్పాపాతులో మా ఇంటి నుండి వెళ్లిపోయాడు.

Tgt2: पार्टी के बाद वह पलक झापकते ही हमारे घर से चला गया।

Explanation:

Tgt1: The term **blink of an eye** has a native term in Telugu రెప్పాపాతులో (read as reppa-paatu-lo).

Tgt2: The term **blink of an eye** has a native term in Hindi **पलक झापकते** (read as palak jha-pa-ka-the).

Example6:

Src: Just add a **pinch of** salt to the rice.

Tgt1: అన్నంలో చిటీకెదు ఉపును వేస్తే చాలు.

Tgt2: चावल में बस एक चुटकी नमक डाल दें।

Explanation:

Tgt1: The term **pinch of** has a native term in Telugu చిటీకెదు (read as chi-ti-ke-du).

Tgt2: The term ***pinch of*** has a native term in Hindi चुटकी(read as chutki).