Flipkart Product Attribute Extraction from Text Challenge

Problem Overview

The hackathon for students is to **leverage a pre-defined data-set from Flipkart to enable product attribute extraction using product description**

Flipkart product attribute extraction challenge consists of **developing a model that extracts product attributes from product description.** For example neck shape, sleeve length, pattern etc are visual attributes of t-shirt. Simply put, given a description of product, the model output attributes of product as shown in figure (refer Table 2).

S. No	Description	Category
1	This T-shirt has a henley neck, Full sleeves with brand logo with comfortable a touch of authenticity. This is made from Cotton can be gone for any casual or special.	T-shirt
2	Off White textured A-line kurta made of pure khadi cotton . While the fabric will take care of your comfort, the simple design with mandarin collar, side and front slits will certainly make you look elegant and dignified	Kurta
3	Blue printed kurta, has a round neck with a full button placket, three-quarter sleeves with roll-up tab feature, curved and vented hem, back hem longer than the front	Kurta
4	Minaro offers this plain formal shirt that is all you need to look your stylish best for a formal look in office. Made from cotton, this design is tailored for a regular fit. Pair this 3/4 sleeve shirt with a pair of black trousers and pumps to complete your ensemble	Shirt

Table 1. Product description with category

You will be provided a sample data set (from multiple categories) which has description & category of product. Your model should take as input a description and produce the product attributes at token/word level. Schema of attributes will be provided which contains all allowed values and description of the attribute.

Your model will be validated against a blind set, which will be shared with you which will contain only the product description.

The performance metric used is precision and recall at attribute level and macro average of it.

You have to provide a solution that outputs tags (at word level) of given product description.

Please note that no supervised data is provided and hence it will be unsupervised problem