

# AI-Driven Mapping of Retail Taxonomies- Part 2

13th Jan, 2021



BY SHWET KAMAL

## Mapping product taxonomies using Deep Learning

In [Part 1](#) we discussed the importance of Retail taxonomy and the applications of mapping retail taxonomies in Assortment Analytics, building Knowledge Graph, etc. Here, we will discuss how we approached the problem of mapping retail taxonomies across sources.

We solved this problem by classifying every retail product to a standard DataWeave defined taxonomy so that products from different websites could be brought at the same level. Once these products are at the same level, mapping taxonomies becomes straightforward.

We've built an AI-based solution that uses state-of-the-art algorithms to predict the correct DataWeave Taxonomy for a product from its textual information like Title, Taxonomy and Description. Our model predicts a standard 4 level (L1-L2-L3-L4) taxonomy for any given product. These Levels denote Category, Sub Category, Parent Product Type and Product Type respectively.

### Approach

Conventional methods for taxonomy prediction are typically based on machine learning classification algorithms. Here, we need to provide textual data and the classifier will predict the entire taxonomy as a class.



We used the classification approach as a baseline, but found a few inherent flaws in this:

- A Classification model cannot understand the semantic relation between input text and output hierarchy. Which means, it cannot understand if there's any relation between the textual input and the text present in the taxonomy. For a classifier, the output class is just a label encoded value
- Since the taxonomy is a tree and each leaf node uniquely defines a path from the root to leaf, the classification algorithms effectively output an existing root-to-leaf path. However, it cannot predict new relationships in the tree structure
- Let's say, our training set has only the records for "Clothing, Shoes & Jewelry > Men > Clothing > Shorts" and "Clothing, Shoes & Jewelry > Baby > Shoes > Boots", Example:

{'title': "Russell Athletic Men's Cotton Baseline Short with Pockets - Black - XXX-Large",

'dw\_taxonomy': "Clothing, Shoes & Jewelry > Men > Clothing > Shorts"}},



Brand: Russell Athletic

Russell Athletic Men's Cotton Baseline Short with Pockets - Black - XXX-Large

★★★★★ 194 ratings

Price: £15.22

Import Fees Deposit Included

Pay £15.22 £0.00: get a £20 Amazon Gift Card on approval for the Amazon Platinum Mastercard. Terms apply.

**Note:** This item is eligible for **FREE click and collect** without a minimum order. [Details](#)

- Relaxed Fit: Most generous, relaxed fit. Extra room for a full spectrum of movement and comfort.
- Stays soft and durable wash after wash
- 50% Cotton, 50% Polyester
- Machine Wash
- Fastening: elastic
- Cotton Baseline Short With Pockets
- Relaxed Fit

✓ [Show More](#)

{'title': "Surprise by Stride Rite Baby Boys Branly Faux-Leather Ankle Boots(Infant/Toddler) - Brown -",

'dw\_taxonomy': "Clothing, Shoes & Jewelry > Baby > Shoes > Boots"}},



Brand: Surprise Stride Rite

## Surprise by Stride Rite Baby Boys' Branly Faux-Leather Ankle Boots (Infant/Toddler)- Brown -

★★★★☆ 1 rating

Size: 3 Toddler

Available from these sellers.

- Hook and Loop closure
- Faux-Leather
- Ankle Length Boots
- Memory Foam Cushion Insole, Rubber Sole for Secure Grip, Hook and Loop Closure for a Comfortable Wear
- Shoe made ideal for: First Walkers, for curious babies and wobbly walkers
- Machine washable, Air Dry

Now, if a product with Title “*Burt’s Bees Baby Baby Boys’ Terry Short*” comes for prediction, then the classifier will never be able to predict the correct taxonomy. Although, it would have seen the data points of Shorts and Baby.



[Visit the Burt's Bees Baby Store](#)

## Burt's Bees Baby Boys' Terry Short

 Climate Pledge Friendly

Price: **\$11.00** + \$12.49 Shipping & Import Fees Deposit to India [Details](#)

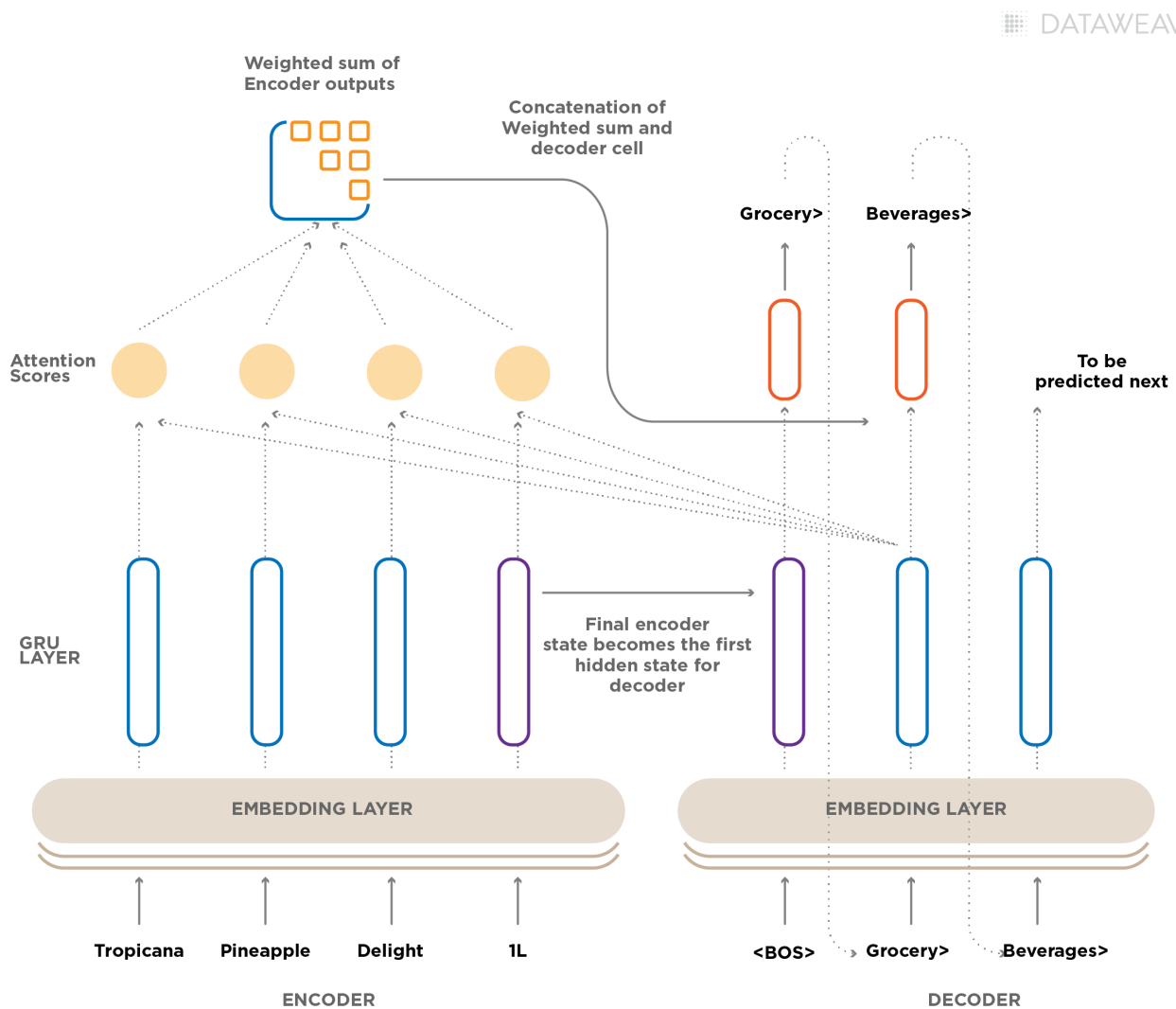
**Size:** 12 Months [Size Chart](#)

**Color:** Olive Sprig

- Pull on

E-commerce product taxonomy has a very long tail, i.e. there's a huge imbalance in counts of data per taxonomy. Classification algorithms do not perform well for very long tail problems.

### Encoder-Decoder with Attention for Taxonomy Classification



## What is Encoder-Decoder?

Encoder-Decoder is a classical Deep Learning architecture where there are two Deep Neural Nets, an Encoder and a Decoder linked with each other to generate desired outputs.

The objective of an Encoder is to encode the required information from the input data and store it in a feature vector. In case of text input, the encoder is mostly an RNN or Transformer based architecture and for image input, it is mostly a CNN-based architecture. Once the encoded feature vector is created, the Decoder uses it to produce the required output. The Encoder and Decoder can be interfaced by another layer which is called Attention. The Role of Attention mechanism is to train the model to selectively focus on useful parts of the input data and hence, learn the alignment between them. This helps the model to cope effectively with long input sentences (when dealing with text) or complex portions of images (when input is an image).



Instead of classification-based approaches, we use an Encoder-Decoder architecture and map the problem of taxonomy classification to the task of machine translation (MT) AKA, Seq2Seq. An MT system takes the text in one language as input and outputs its translation as a sequence of words in another language. In our case, the input maps to the textual description of a product, and the output maps to the sequence of categories and sub-categories in our taxonomy (e.g., Clothing, Shoes & Jewelry > Baby > Shoes > Boots). By framing taxonomy classification as an MT problem, we overcome a lot of limitations present in classical classification approaches.

- This architecture has the capability to predict a taxonomy that is not even present in the training data.
- Talking about the example we discussed earlier where a traditional classification model was not able to predict the taxonomy for “*Baby Boys knit terry shorts – cat & jack gray 12 m*”, this Encoder-decoder model easily predicts the correct taxonomy as “*Clothing, Shoes & Jewelry > Baby > Clothing > Shorts*”
- We achieved a much higher accuracy because the model understands the semantic relationship between the input and output text, as well as giving attention to the most relevant parts in the input, when generating the output

input = south of france lavender fields bar soap  
 output = Beauty\_&\_Personal\_Care > Skin\_Care > Body > Soaps <EOS>

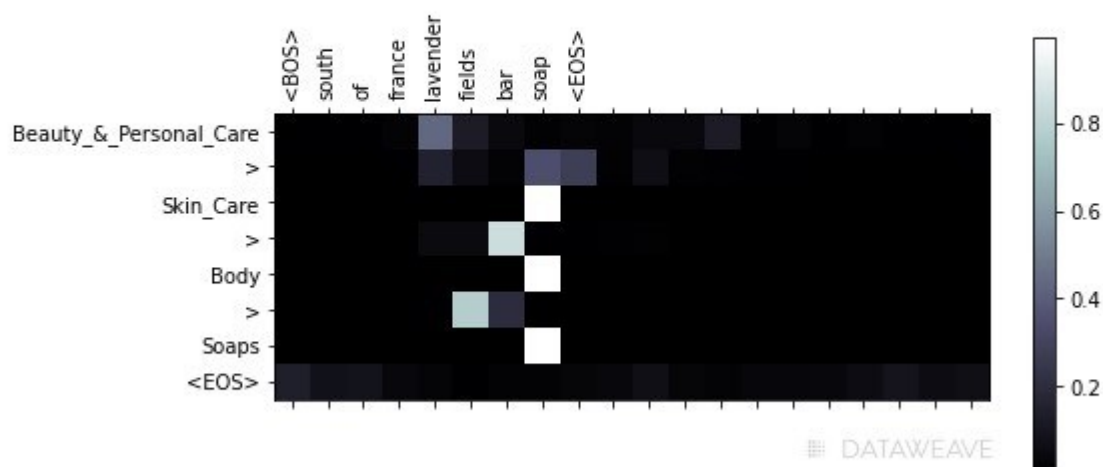


Fig. Attention visualization for product title “**South of France lavender fields Bar Soap**”. It can be seen from the image that the attention weights of “soap” word is very high when predicting the output at different time-steps.

We used pre-trained fasttext word embeddings to vectorize textual input, pass on to the GRU-RNN based encoder which processes the input sequentially, and generates the final encoded vector. The Decoder which is also a GRU-RNN takes this encoded input and generates the output sequentially. Along with the encoded vector, there is also an attention vector which is passed to the Decoder for the output at every time-step.

We trained both the Classification model (Baseline) and the Encoder-Decoder model for the Fashion category and the Beauty & Personal Care category.

For Fashion, we trained the model with 170,000 data points and validated it on a 30k set. For Beauty Category, we trained the model on 88k data points and validated it on a 20k set. We were able to achieve 92% Seq2Seq accuracy in 1,240 classes for the Fashion category and 96% Seq2Seq accuracy in 343 classes for the Beauty Category, using the Encoder-Decoder approach.

## Summary and the Way Forward

Since we moved to this approach, we have seen drastic improvements in the accuracy of our Assortment Intelligence accounts. But the road doesn't end here. There are several challenges to be tackled and worked upon. We're planning on making this process language agnostic by using cross-lingual embeddings, merging models from different categories and also using product Image to complement the text-based model with visual input via a Multi-Modal approach.

### References

<https://arxiv.org/pdf/1812.05774.pdf>

<https://sigir-ecom.github.io/ecom2018/data-task.html>

<https://arxiv.org/pdf/1703.03906.pdf>

- **Shwet Kamal Mishra**

13th Jan, 2021

ARTIFICIAL INTELLIGENCE

E COMMERCE