**[Fynd](javascript:void(0)" \t "_self)** is the world’s first inventory less omnichannel e-commerce company. Since we directly get the content catalogue from brand’s stores, we have to validate and often classify many attributes of a product as per the Fynd content guidelines. In Fynd, we use machine learning for various tasks such as dynamic pricing, profile-based feed, product recommendation, and automated product attribute tagging etc. Our work is nicely summarized in the [**presentation**](javascript:void(0)). We are looking for an intern who will be working with the core ML team to solve similar problems.

**Data Description**

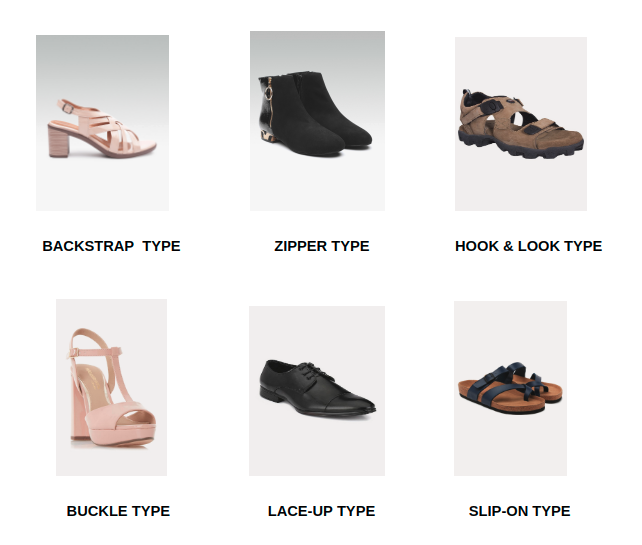
Public dataset: Contains a CSV file with a column name ‘class’ and image URLs to different views. Download images from URL for respective classes. Download the dataset.

Note: We have 20% private dataset for our own evaluation.

**Problem Statement**

1) Each product has multiple views (front, left, right, back) and only a few views have the information about the product through which you can make the ML model learn its features. Build ML model and explain your approach on how will you select the specific view from the given set of different views.

Example: In the case of footwear, Side view(left/right) has most information regarding the type of heel (high, medium or flat).Similarly, different views have different information for each use case.  
(Flat type)



                                                 FOOTWEAR: CLOSURE TYPES

2) Identify the closure type of footwear from the data (CSV) provided which has imbalanced data distribution across 6 classes. You can collect more data using web scraping or different sources. Build ML model which effectively identifies the closure type of the footwear.   
Here is a link to the visual document to understand each class of the data.

**Submission**

* Make a final python script which takes a CSV as an input with a column [image\_name] and gives prediction across each image\_name on a column named as ‘label’ (Preferably use argument parser in the script)
* Model weights and definition files
* All training and inference related codes (either zip or repo URL)

Make a PPT explaining your approach towards the problem. Specify key points such as:

* Provide your training, validation accuracy and confusion matrix on the public dataset. Explain other evaluation metrics which can be used
* Did you feel the need for more data? If yes, explain different ways you collected the data
* Explain all the hyperparameters used and how did you end up using them only
* (Optional) Key insights about the public dataset given. Any special observation
* What else can you do for better performance?
* Reference to all the research papers/Github issues/blogs etc which helped you in solving the problem
* Submit your code/repo at **code@gofynd.com**with your latest resume
* **We don’t have a time limit for the test. We recruit in FIFO fashion.**
* You can use google colab for training your models.

**Reward**

* Internship in Fynd with a stipend as a part of the core ML team which can be further extended
* NVIDIA GPU

**Links**

* [**dataset -> https://docs.google.com/spreadsheets/d/1r0k3IHvvLZLtdOVAQwt2dwQofuC5uS\_idL-sLZ3jw5c/edit#gid=806989492**](javascript:void(0))
* [**what we do -> https://docs.google.com/presentation/d/1X7x6fofiQ4TLecBEYLgrXRD5O19-VfLOaP68Pwwv8n8/edit#slide=id.gd6b24c821\_0\_4**](javascript:void(0))
* [**https://www.fynd.com/dune-london-genovia-loafers-431317-d6fa84.html#source=search&query=footwear&pageno=1&position=1&index=1&source\_id=431317&sub\_source=product\_grid**](javascript:void(0))
* [**https://blog.gofynd.com/machine-learning-internship-challenge-2019-6b4e9dddb637**](javascript:void(0))