Apparel & Footwear Prediction

Team: A4
Names: Becky Wang, Chia-Chien Chang, Xinying Wu, Vy Nguyen, Yunling Li

Footwear

Apparel

TABLE OF CONTENTS

Footwear

Apparel

01

Introduction

Introduction

Outcome variable: product category

- Binary : 0 = Apparel; 1 = Footwear

Business Application: automated product categorization for e-commerce

- Reduce manual tagging and errors in large product catalogs
- Improve search and filtering, making it easier for shoppers to find products
- Enhance recommendations and visual search, leading to better user experiences
- Detect miscategorized listings, keeping platforms well-organized

02

Performance

Footwear

Apparel

Performance

Metric(s)

We primarily use "**accuracy**" as our evaluation metric:

- The classes are reasonably balanced (Apparel = 46%, Footwear = 54%), so accuracy is meaningful
- It's an intuitive metric for business users to understand and aligns with the goal of **correct classification**

Human Performance

Approximate human performance for classifying Apparel vs. Footwear is close to **100%**, as the distinction is visually obvious to humans.

03

Dataset

Dataset Overview

Total image

Total images: 2906

Apparel(label 0)

Apparel: 1326

Ratio: 46%

Footwear(label 1)

Apparel: 1580

Ratio: 54%

Image Sizes & Examples

Image sizes **vary**, but common dimensions are

(1800, 2400) and **(1080, 1440)**

Data Split

01

Training set

1,743 images

02

Validation set

581 images

03

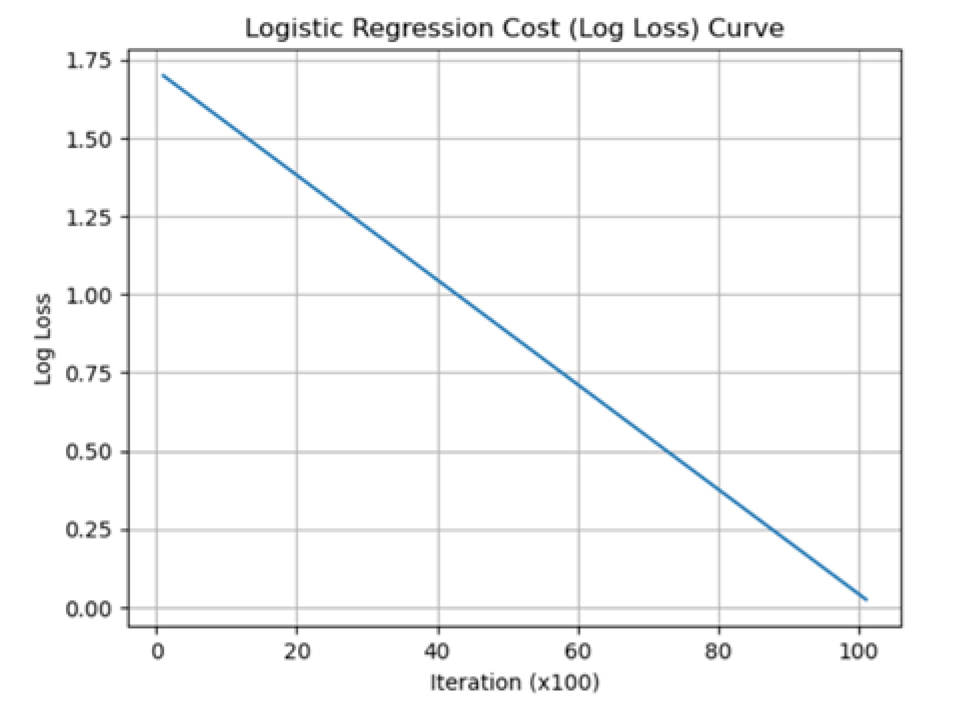
Test set

582 images

The dataset was split using stratified sampling to preserve class balance, following a **60%/20%/20% ratio**.

04 Base Model

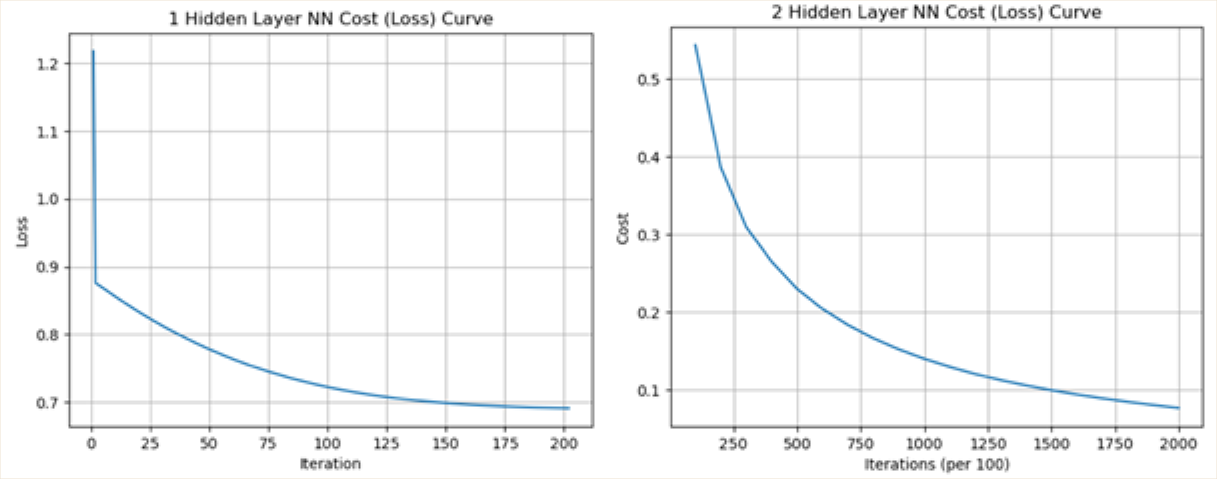
Logistic Regression (no hidden layers)



Learning rate: 0.001
Training Accuracy: 99.7%
Validation Accuracy: 99.3%

- No signs of overfitting or underfitting
- Loss is smoothly decreasing, suggesting well-tuned learning rate and parameters

Neural Network

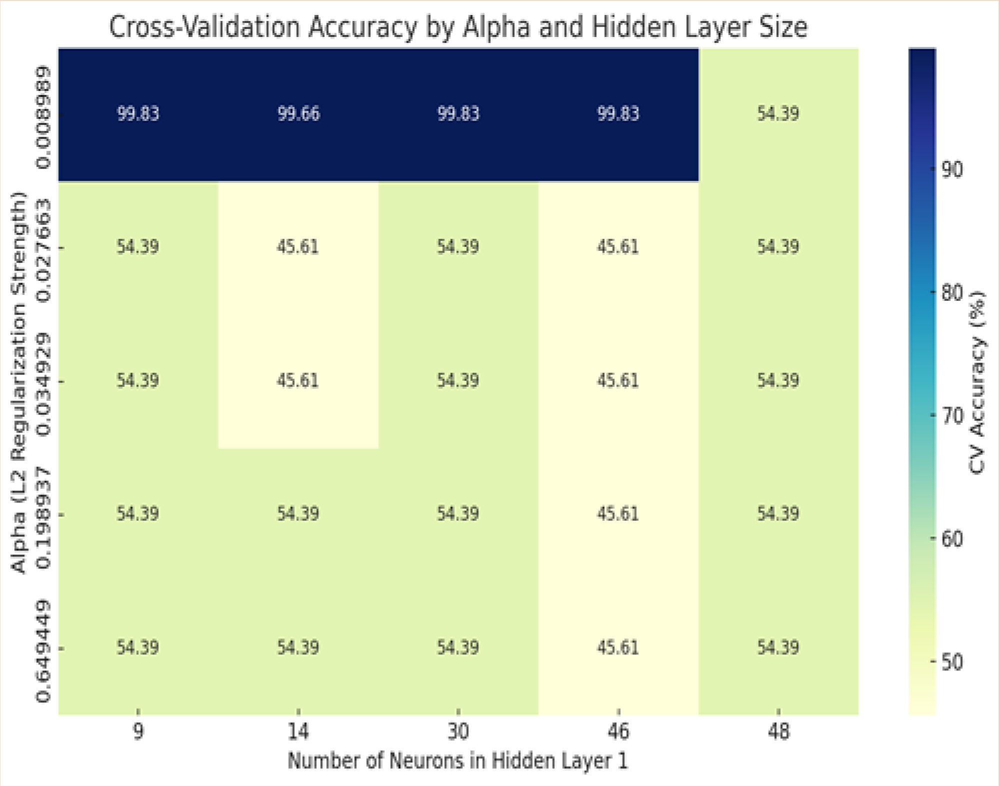


- Performance increased with a bigger neural network
- Loss graph of 2 hidden layers NN shows a more expressive model that captures patterns more effectively
- No signs of overfitting or underfitting

	Training Accuracy	**Cross-validation Accuracy**
1 hidden layer of 4 units	54.3%	54.3%
2 hidden layers (7 units - 4 units)	99.5%	100%

05 Other Models

5.1 Performance of 5 different learning rates



Best learning rate: 0.0089
Best hidden unit for this learning rate: 30

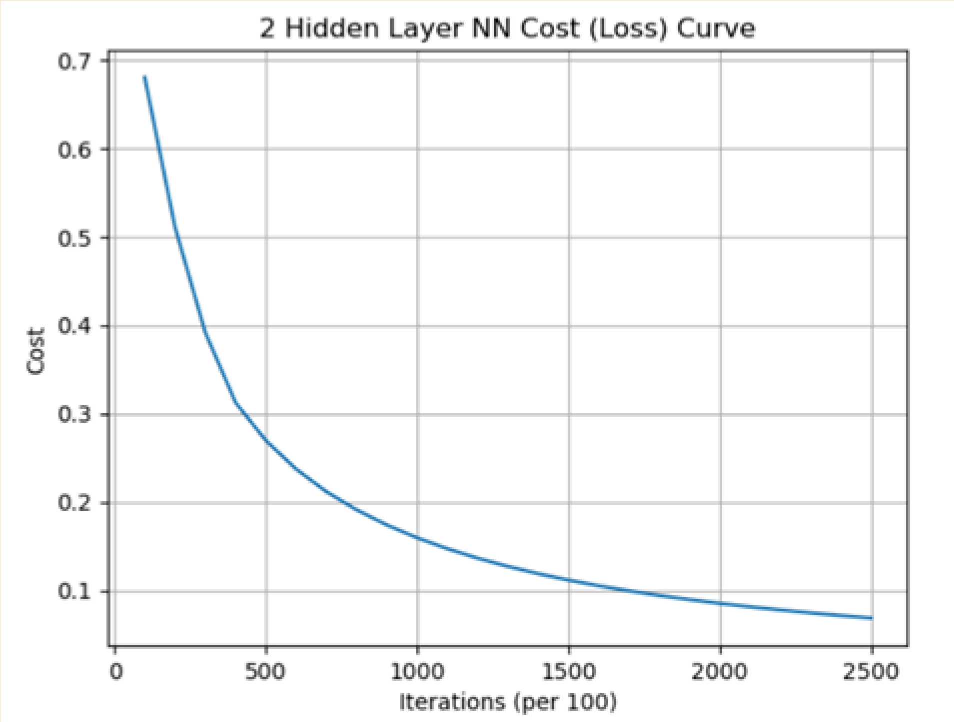
Training Accuracy: 99.6%
Cross-validation Accuracy: 99.8%

- No clear improvement
- Cross-validation Accuracy dropped but closer to the training accuracy
- No signs of overfitting or underfitting

5.1 Performance of 5 different learning rates

New model architecture
Learning rate = 0.0089
Hidden layer 1: 30, ReLU activation
Hidden layer 2: 4, ReLU activation

- Curve is smooth and progressively flattens out
- Loss is consistently decreasing



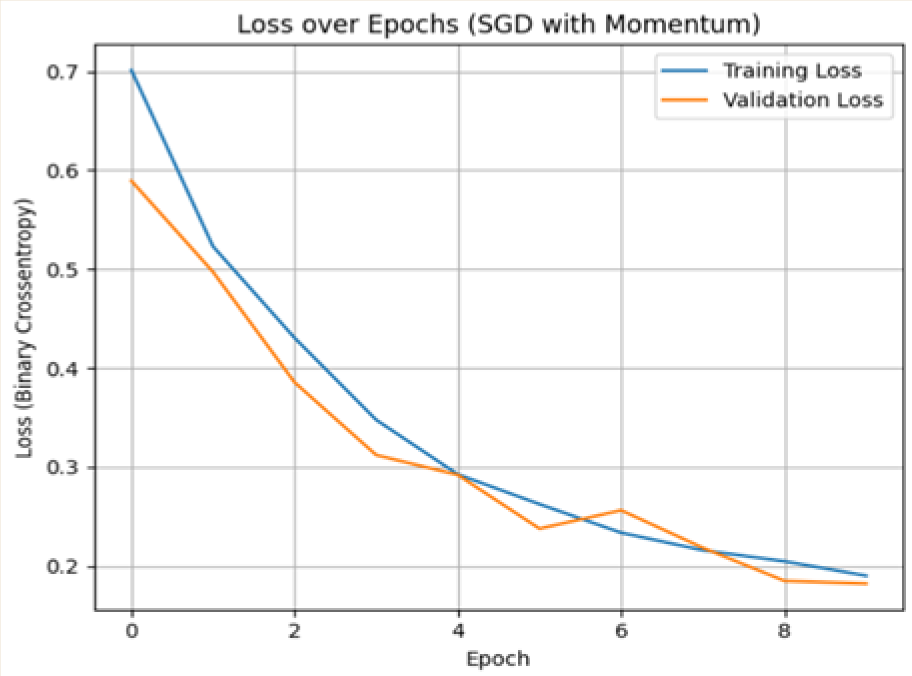
5.2 Performance of different optimization algorithms

Best algorithm: SGD with Momentum = 0.5

Training Accuracy: 98.7%
Cross-validation Accuracy: 98.6%

- No sign of overfitting or underfitting
- Both training & validation loss steadily reduce

Optimizer	Train Loss	Train Accuracy
Adam	0.729587	0.543890
SGD	0.268450	0.970166
Momentum	0.189582	0.986804
RMSprop	0.700199	0.543890



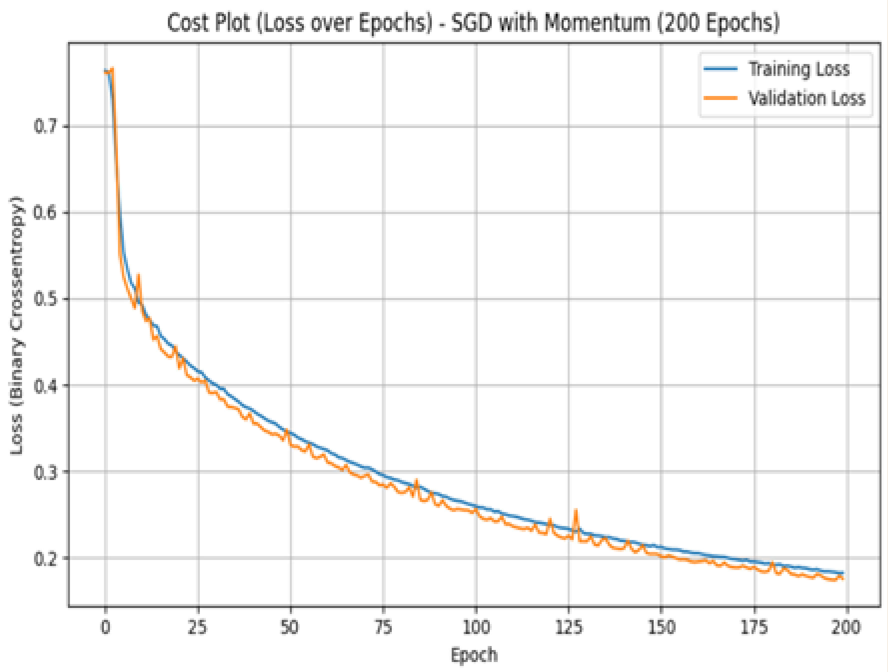
5.3 Performance of different numbers of epochs

Best epochs for SGD with Momentum: 200 epochs

Training Accuracy: 99.6%
Cross-validation Accuracy: 99.8%

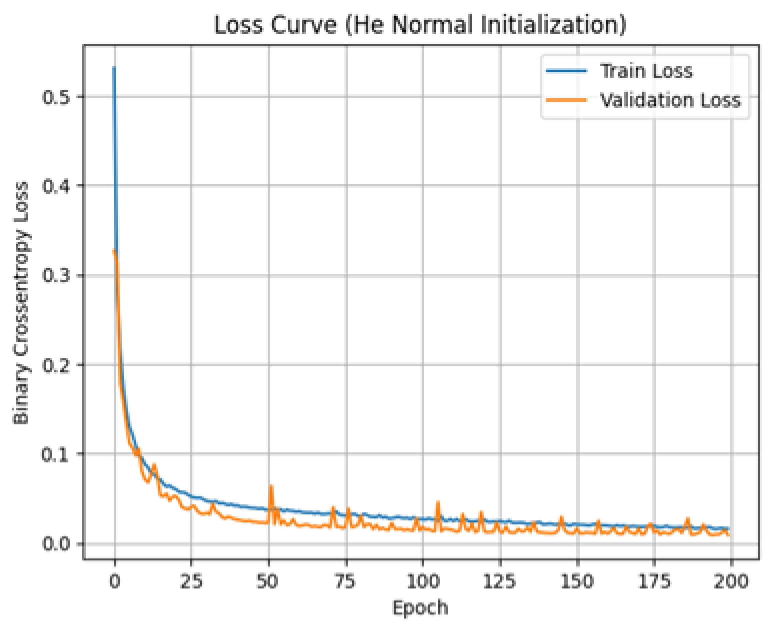
- Accuracy increased 1% from the previous model
- No sign of underfitting or overfitting
- Both training & validation loss steadily reduce

Epochs	Train Loss	Train Accuracy
50	0.757383	0.543890
100	0.248084	0.991394
200	0.181778	0.995984



5.4 Performance of different Weight Initialization

Initialization	Train Accuracy	Validation Accuracy
TF Default	0.9920	0.9983
He Normal	0.9943	1.0000
Random Normal (stddev=0.05)	0.9908	1.0000
He Uniform	0.9925	1.0000

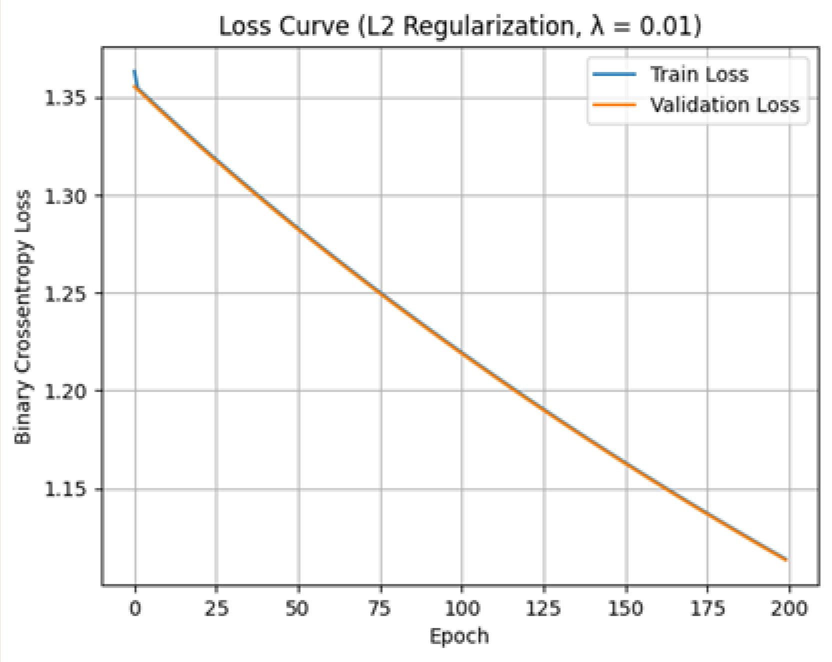


Best weight initialization: **He Normal**

- No clear improvement over baseline
- Fast drop within the first 25 epochs
- Both train & validation loss decrease smoothly
- No major train-val gap → strong generalization
- No sign of overfitting

5.5 Performance of different L2 Regularization

Lambda (λ)	Train Accuracy	Validation Accuracy
0.000	0.9937	0.9983
0.001	0.9925	0.9948
0.010	0.9931	1.0000
0.100	0.9862	0.9948

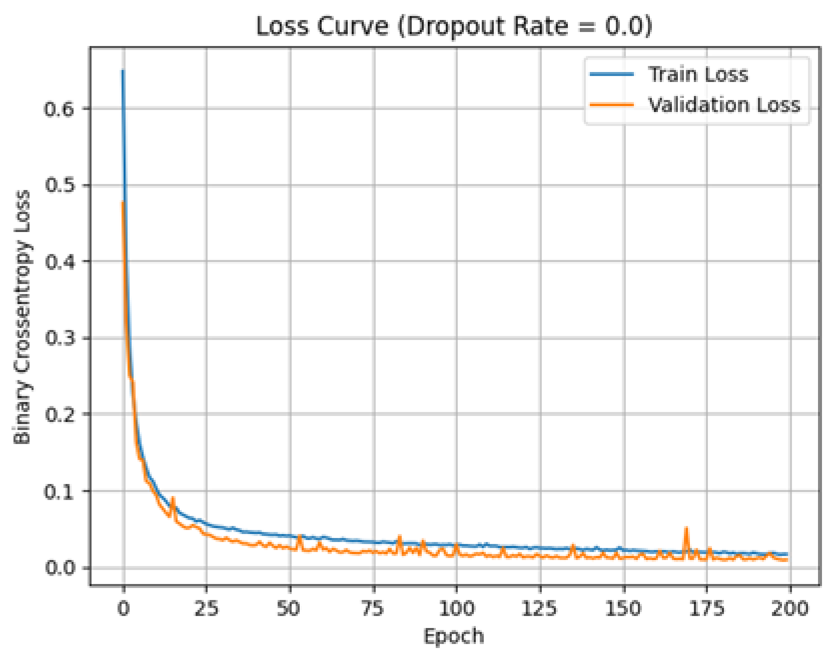


Best L2 Penalty: λ = 0.01

- No clear improvement over baseline
- Slower but steady reduction in loss
- Both train & validation loss closely align
- No major train-val gap → strong generalization
- No sign of overfitting

5.6 Performance of different Dropout Rates

Dropout Rate	Train Accuracy	Validation Accuracy
0.0	0.9960	0.9966
0.2	0.9925	1.0000
0.4	0.9954	0.9966
0.6	0.9925	1.0000



Best Dropout: **rate = 0.2 or 0.6**

- No clear improvement over baseline
- Fast drop within the first 25 epochs
- Both train & validation loss closely align
- No major train-val gap → strong generalization
- No sign of overfitting

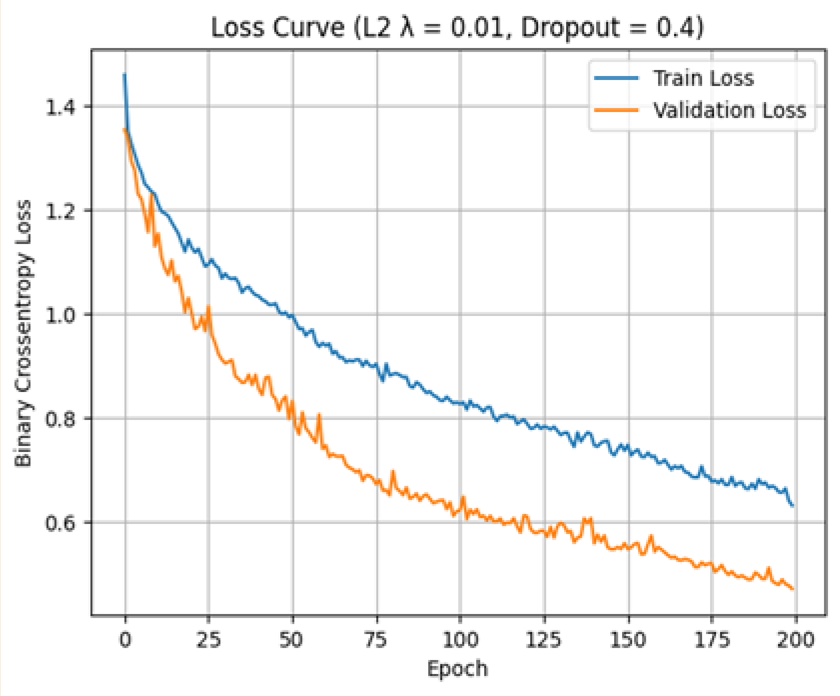
5.7 Performance of different Mixtures of L2 & Dropout

L2 λ (0, 0.01, 0.1)	Dropout Rate (0, 0.4, 0.6)	Train Accuracy	Validation Accuracy
0.00	0.0	0.9971	0.9983
0.00	0.4	0.9908	0.9966
0.00	0.6	0.9897	0.9914
0.01	0.0	0.9948	0.9966
0.01	0.4	0.9937	1.0000
0.01	0.6	0.9931	1.0000
0.10	0.0	0.9925	1.0000
0.10	0.4	0.9920	0.9983
0.10	0.6	0.9920	0.9966

5.7 Performance of different Mixtures of L2 & Dropout

Best Combination: λ = 0.01 & Dropout = 0.4
Train accuracy = 99.37%
Validation accuracy = 100%

- The mixture did not outperform L2 or dropout regularization alone
- Train loss decreased more slowly than validation → reduce overfitting
- Validation loss always < train loss -> strong generalization
- Stable learning without performance degradation
- No sign of overfitting



5.8 Performance for Different Batch-normalizations

Model	Specification	Training Accuracy	Validation Accuracy
No BN	No Batch-normalization layers (baseline model)	0.9960	0.9983
BN 1	Default Batch-normalization (after each Dense hidden layer)	0.9983	0.9966
BN 2	Custom Batch-normalization with momentum=0.9, epsilon=1e-4 after each layer	0.9994	0.9983

5.9 - 5.12 Advanced Modeling Techniques

Early Stopping

Did not improve accuracy,
but it helps prevent overfitting

Training Accuracy: 99.31%
Validation Accuracy: 100.00%
Model Selection:
Use NN model as an example

CNN Model

Overfits quickly

Training Accuracy: 99.89%
Validation Accuracy: 99.66%
Model Selection:
Conv1, Pool1, Conv2, Pool2 layers, and a flattened layer

Pre-trained Model

Fast convergence,
perfect validation

Training Accuracy: 99.94%
Validation Accuracy: 100.00%
Model Selection:
MobileNetV2

RNN Model

Close to NN/CNN

Training Accuracy: 98.80%
Validation Accuracy: 99.66%
Model Selection:
GRU

Footwear

Apparel

06

Best Model

Final Model Selection

Final Model Architecture

Input → Hidden Layer 1 (30 neurons, ReLU) → Hidden Layer 2 (4 neurons, ReLU) → Output (Sigmoid)

- SGD with Momentum = 0.5
- Learning Rate: 0.001
- L2 Regularization (λ): 0.001
- Weight Initialization: He Normal
- Epochs: 200

Model Type	Train Accuracy	Validation Accuracy
DNN (30-4)	99.6%	99.83%
Logistic Regression	99.7%	99.3%
DNN (7-4)	99.5%	99.83%

Test Accuracy for DNN 30-4: 99.83%

07

Managerial Implications

Footwear

Apparel

Deep Learning vs. Logistic Regression

Performance Comparison

- Logistic Regression Validation Accuracy: 99.3%
- DNN Model (30-4) Validation Accuracy: 99.83%

+ 0.53% Accuracy Improvement

- A small improvement can bring a significant business impact at scale.

Why This Matters in Practice

- Fewer misclassified items → Better search & recommendations.
- More accurate tagging → Reduced manual rework.
- Less user frustration → Higher conversion rate & trust.
- Improved scalability for future product expansion.

From Model Accuracy to Business Value

Business Context

Our platform processes 10,000+ new product images daily. Each classification error has downstream impacts across:

- Manual correction labor
- Customer service load
- Misdirected recommendations
- User frustration & churn risk

Quantified Impact

Model	Error Rate	Errors/Day	Cost/Day
Logistic Regression	0.7%	70	$140
Final DNN (30-4)	0.17%	17	$34
Daily Cost Saving	—	↓ 53	**$106**
Monthly Saving	—	—	**~$3,180**

- *Assumes $2 per misclassification (labor, rework, lost revenue, etc.)*

Beyond the Numbers: Strategic Benefits

- Scalability - Handles growth without adding headcount.
- Customer Experience & Trust - Better data → easier discovery → higher conversion.
- Competitive Advantage - Faster category expansion and market response.

THANK You !

Footwear

Apparel