

JUDGING A BOOK BY ITS COVER

Analyzing Book Covers and Titles for Bestseller
Prediction with Transfer Learning

Angel Gantzia

Nick Antoniou

Dido Stoikou

Pavena Vongkhammi



PROBLEM OVERVIEW



Listopia > Book Covers Book Lists



Judge A Book By Its Cover!

12,531 books — 13,107 voters



Best Book Cover Art

11,541 books — 7,150 voters



Beautiful Book Covers of 2011

727 books — 4,284 voters



Best Book Covers for 2010

357 books — 3,189 voters



Cover Lust

919 books — 564 voters



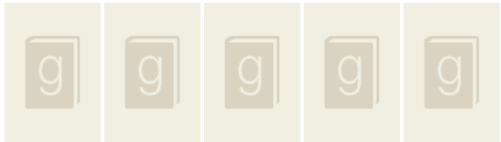
Haunting and Spine-Chilling Covers

343 books — 265 voters



Awesome Covers that DON'T Have Shirtless Guys or Girls in Dresses

740 books — 256 voters



CAN'T WAIT TO SEE THE COVER!!!

369 books — 2,097 voters

PROBLEM OVERVIEW



Listopia > Book Covers Book Lists



Judge A Book By Its Cover!

12,531 books — 13,107 voters



Best Book Cover Art

11,541 books — 7,150 voters



Beautiful Book Covers of 2011

727 books — 4,284 voters



Best Book Covers for 2010

357 books — 3,189 voters



Cover Lust

919 books — 564 voters



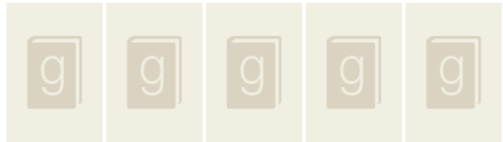
Haunting and Spine-Chilling Covers

343 books — 265 voters



Awesome Covers that DON'T Have Shirtless Guys or Girls in Dresses

740 books — 256 voters



CAN'T WAIT TO SEE THE COVER!!!

369 books — 2,097 voters

Clearly, readers find value in beautiful covers



Can book cover designs directly influence consumer purchases (impulsive buys) on platforms like Amazon?



We explore whether the **cover design** and optionally its **title** are **sufficient predictors** of its chance to become a **best-seller**?

WHY DOES THIS MATTER?



Data Driven Decisions for **Publishers**

Publishers gain consumer intelligence and take data driven decisions based on predictive models, guided by consumer trends



Support for **Independent Authors**

Independent (even debuting) authors can have a gauge for the visual of their books



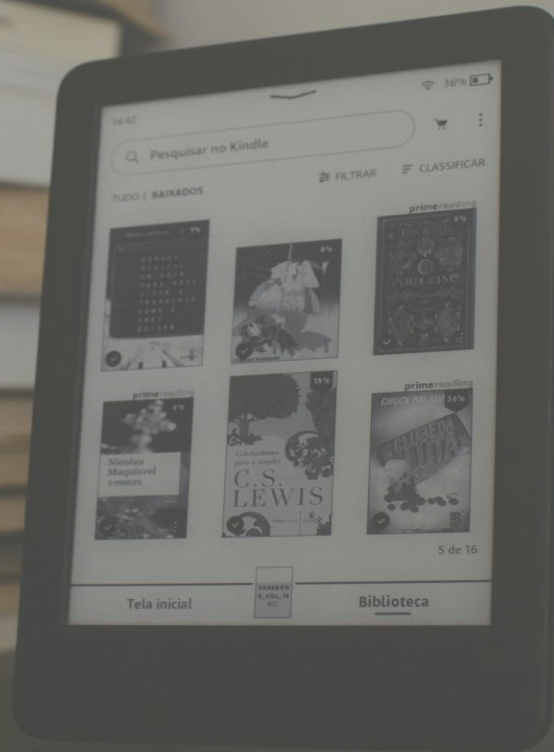
Recommendation/Placement Optimization **for retailers**

Online retailers, like Amazon, can use the findings to improve their recommendation engine and cold start product placement for more engaging and effective shopping



Increasingly beautiful covers for **consumers**

Consumers will be happier with the increasingly more beautiful design covers that their books will have





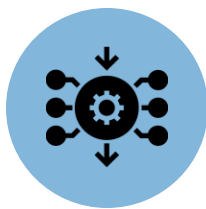
Scrape kindle best-seller list from Amazon

- Title
- ASIN
- Books' cover
- CAPTCHA:** Roadblock for book descriptions



Augment dataset with additional Kaggle data

- ASIN code
- Title
- link to their cover online
- Product page on Amazon



Filter only 7 genres and balance the label

- Mystery, Thriller & Suspense
- Children's eBooks
- Teen & Young Adult
- LGBTQ+ eBooks
- Literature & Fiction
- Science Fiction & Fantasy
- Romance



Scrape the image for balanced Kaggle dataset

Final dataset:
~3600 images of kindle bestsellers and non-bestsellers



Preprocessing the scraped images

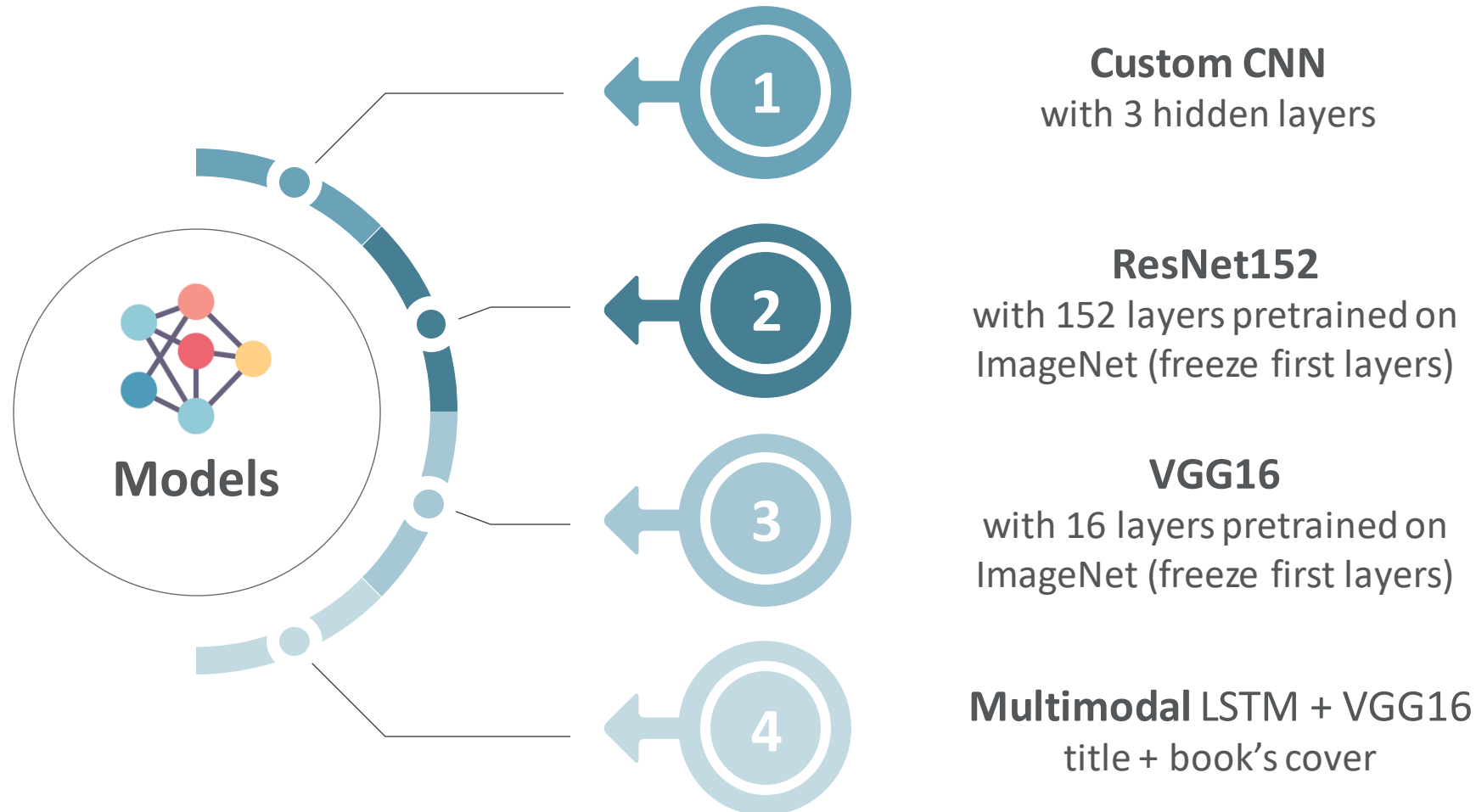
- Normalizing and resizing
- Experimented with cropping and flipping the images



APPROACH



Fine-tuned with 5 hyperparameters; batch size, number of epochs, learning rate, weight decay, and criterion using random search for these 4 models

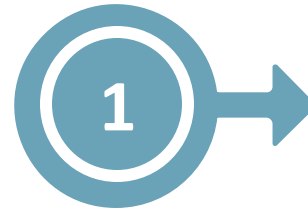


RESULTS



Best Model	Best Hyperparameters	Test Accuracy
Multimodal	Batch size: 64 Number of epochs: 50 Learning Rate: 0.001 Weight Decay: - Criterion: Binary Crossentropy	80.81%
VGG16	Batch size: 64 Number of epochs: 15 Learning Rate: 0.0008 Weight Decay: 0.0001 Criterion: CrossEntropyLoss	80.19%
ResNet152	Batch size: 32 Number of epochs: 10 Learning Rate: 0.0007 Weight Decay: 0.001 Criterion: CrossEntropyLoss	79.87%
Custom CNN	Batch size: 128 Number of epochs: 10 Learning Rate: 0.0004 Weight Decay: 0.001 Criterion: CrossEntropyLoss	78.62%

LESSONS LEARNT



Consumers are **influenced by visual appeal of books** when shopping online



Buyers also buy books on Kindle devices, investigate customer behaviour with **greyscale images**



Book covers are **meaningful predictors** book titles don't improve accuracy, need to **explore book descriptions**

THANK YOU

Do you have any questions?

