dermcheck.ai

final presentation | week 15

Ruby Han, George Jiang, Gerrit Lensink, Shivani Sharma

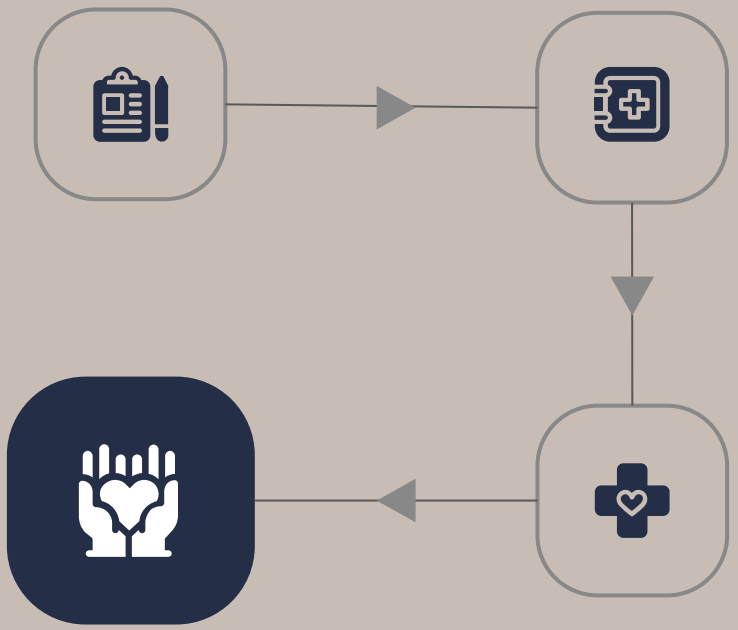
team

Special thanks to Puya Vahabi & Alberto Todeschini
Capstone Instructors, UC Berkeley MIDS Fall 2022

problem

Doctor's Appointments

Are costly in terms of time and money



Patients

should not have to wait to get an idea of their possible diagnosis

because

some skin conditions are more serious than others

dermcheck.ai

can help

impact

1 in 4

Americans are affected by a skin condition at any given time

$15 billion

Expected size of global skin cancer treatment market by 2027

85%

of Americans have a smartphone

Clinicians

Clinicians can use dermcheck.ai to check their diagnosis with AI-backed inference

Patients

Patients can check their possible diagnosis before seeing a doctor, to understand if their condition is serious or not

dermcheck.ai

with your photo, dermcheck.ai can classify up to 5 categories of skin conditions
encompassing ~120 different skin abnormalities
ranging from the most malignant skin cancers to the most common skin diseases

product

User Outcomes

Benign	Moles Nevi Dermatofibroma
Skin Condition	Eczema Acne Dermatitis
Infection	Warts Fungus Ringworm
Malignant	Melanoma Carcinoma Keratosis
Unclassified	

product

User Outcomes

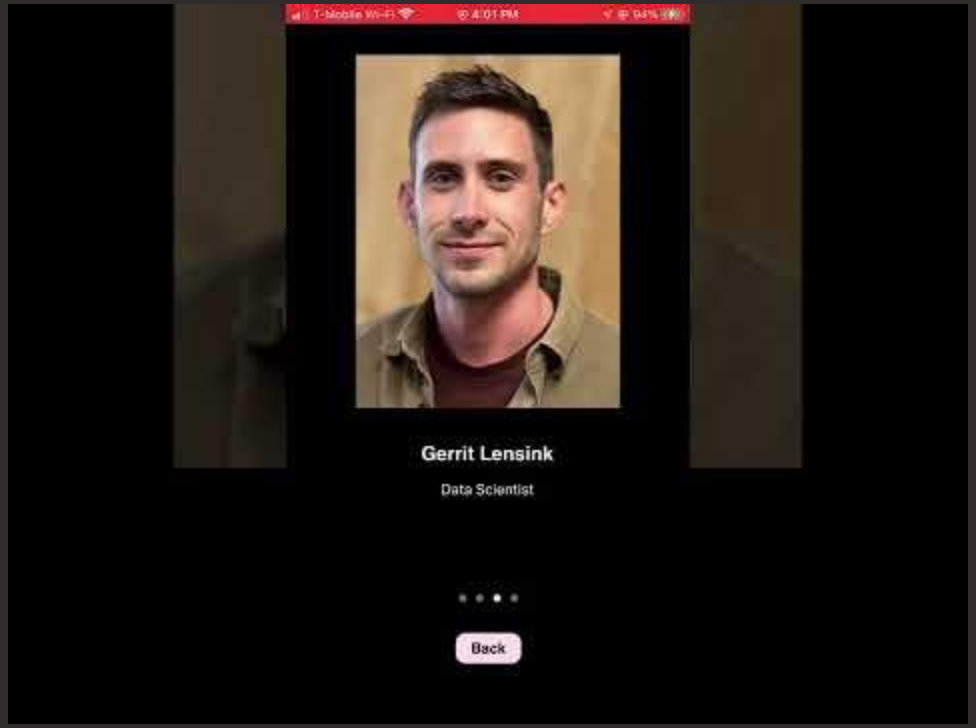
Benign	No Action Required
Skin Condition	Contact Physician (lower risk)
Infection	
Malignant	Contact Physician (higher risk)
Unclassified	

product

Benign	Moles Nevi Dermatofibroma	No Action Required
Skin Condition	Eczema Acne Dermatitis	Contact Physician (lower risk)
Infection	Warts Fungus Ringworm	
Malignant	Melanoma Carcinoma Keratosis	Contact Physician (higher risk)
Unclassified		

product

App Demo



data

ISIC

The International Skin Imaging Collaboration (ISIC) is an academia and industry partnership designed to facilitate the application of skin imaging to help reduce melanoma mortality. ISIC has a large and open source skin image library, creating resources for dermatology and CS communities. We have leveraged data from the ISIC 2018, 2019, and 2020 datasets.

58k images

Dermnet

The Dermnet datasets contains images of around 23 types of skin diseases. Dermnet is the largest dermatology online source built for the purposes of providing online medical education. Dermnet helps thousands of people make informed, evidence-based decision on how to care for skin conditions by providing reliable information.

20k images

Stanford Diverse

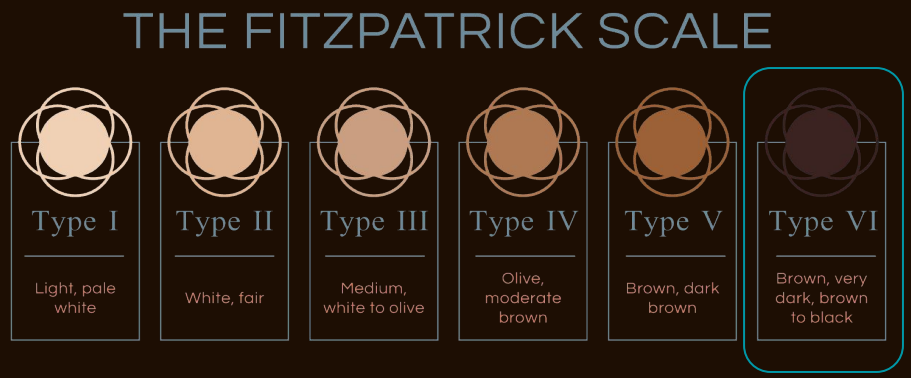
To ascertain potential biases in algorithm performance in this skin disease detection, the Diverse Dermatology Images (DDI) dataset was curated. It is the first publicly available, deeply curated, and pathologically confirmed image dataset with diverse skin tones. The DDI was retrospectively selected from reviewing pathology reports in Stanford Clinics from 2010-2020.

600 images

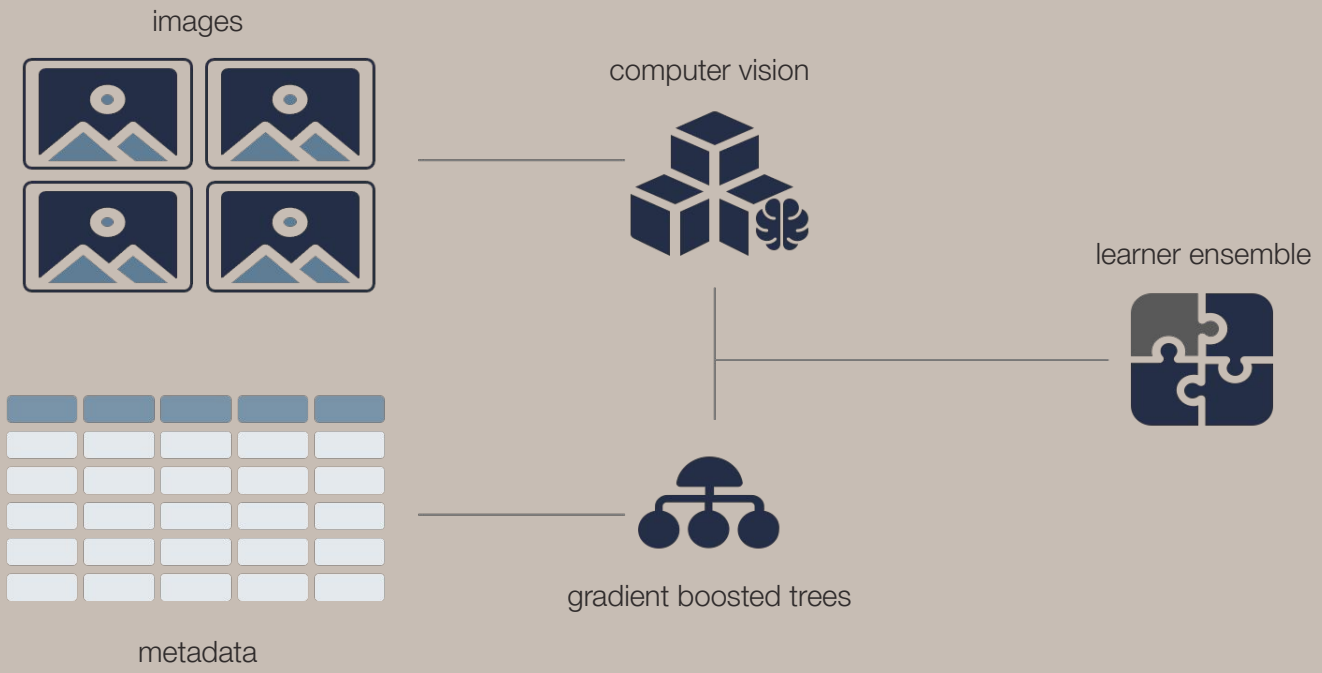
data

Most of the readily available skin disease image data out there is heavily biased towards lighter skin tones. In order to combat this we have used image blending by overlaying two images over one another with the goal of darkening the skin tones in our dataset.

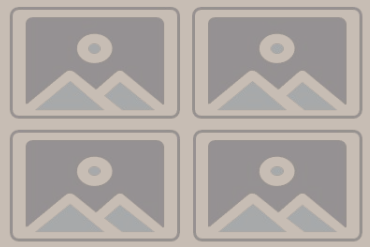
We have leveraged the darkest classification in the Fitzpatrick classification of skin phenotypes in order to do this, and 33% of the images in our final dataset are of this classification.



model

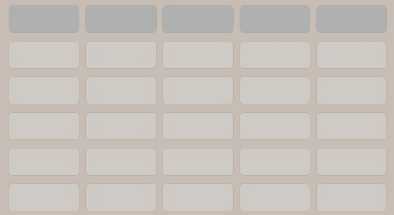

images
computer vision
learner ensemble
metadata
gradient boosted trees

model



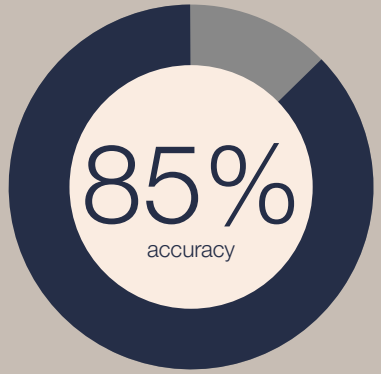
images

computer vision



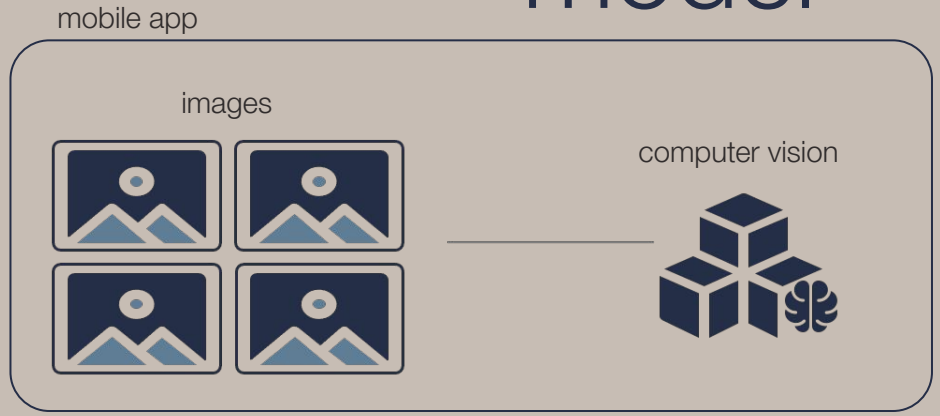
metadata

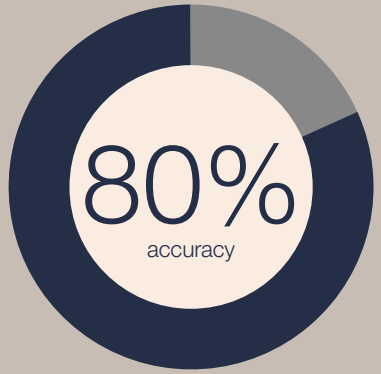
gradient boosted trees



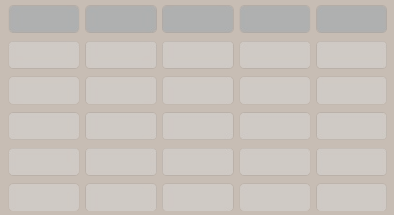
learner ensemble

model





learner ensemble



metadata

gradient boosted trees

Image model

computer vision

Architecture*

Framework: **Pytorch**
Pretrained Models: **Resnet, VGG**
Epochs: **10, 15, 20, 50**
Learning Rate: **.000559, .0025, .003, .0035,**
Optimizer: **SGD, Adam**
Batch Size: **64**
Workers: **24**
Code Base:

* final learner model spans 9 different CV models. Attributes above reflect range of attributes over final models

Metadata model

Decision Tree

Architecture

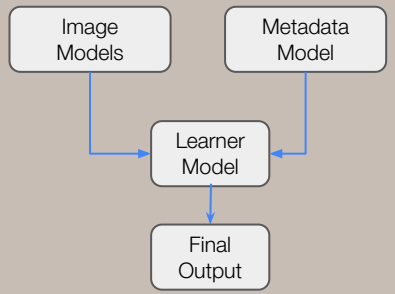
Framework: gradient boost
Tree Depth: 2
Min_child_weight: 0.81
Learning Rate: 0.025
Gamma: .10
Subsample: 0.80
Colsample_by_tree: 0.42

Key Inputs/Learning

Color of the image (Most Important)
Gender
Age
Location of the Image (Least Important)

Ensemble Models

Stacking Ensemble



Architecture

Ensemble Framework:

- Stacking

Learner Model Framework:

- Gradient Boost

# of Input Models: 9

Learner Model Hyperparameter:

- Tree Depth: 1
- Min_child_weight: 0.16
- Learning Rate: 0.033
- Gamma: .25
- Subsample: 0.80
- Colsample_by_tree: 0.40

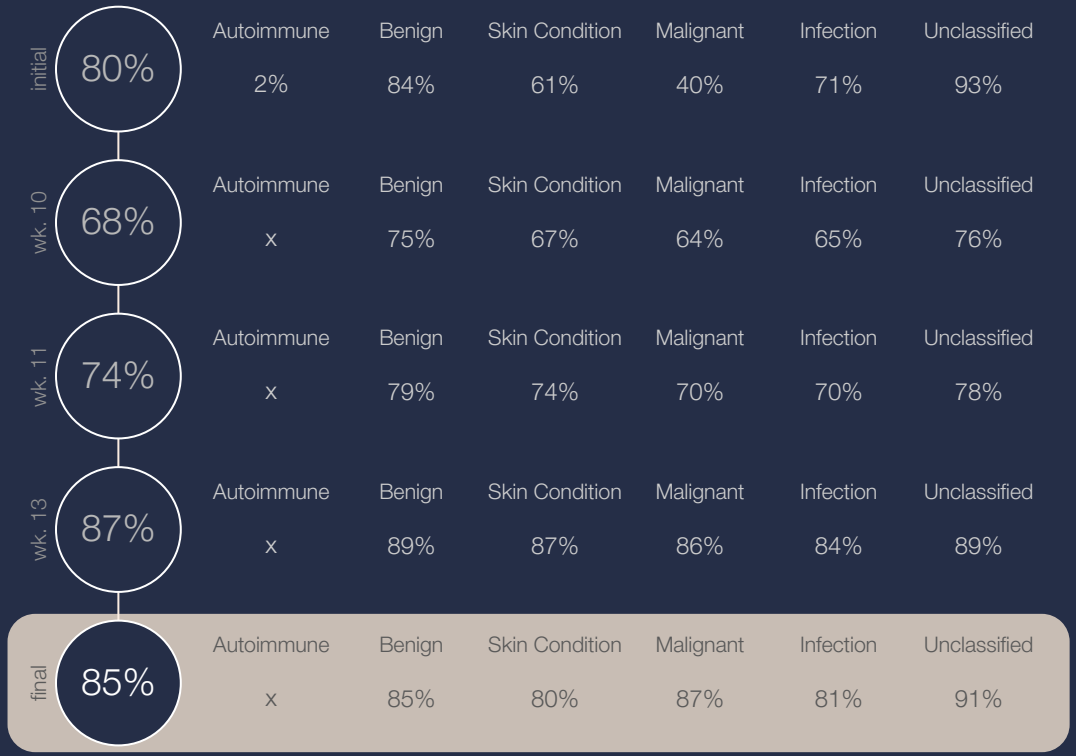
Key Learning

Ensemble result depends on several factors:

- # of models stacked
- accuracy of the models
- how different the models are

model

Improvement Walk



model

initial 80%

Autoimmune	Benign	Skin Condition	Malignant	Infection	Unclassified
2%	84%	61%	40%	71%	93%

Improvements

- Design data strategy
- Design model strategy
- Organize, transform, and clean data
- Implement resnet for baseline models

model

Improvements

- Research confused subclasses
- Iterate over different class and split combinations
- Undersampling
- Remove ambiguously-labeled images from train
- Remove underrepresented images in subclass
- Image Preprocessing

	Autoimmune	Benign	Skin Condition	Malignant	Infection	Unclassified
initial	2%	84%	61%	40%	71%	93%
wk. 10	x	75%	67%	64%	65%	76%

model

Improvements

Gridsearch
- LR & transforms only

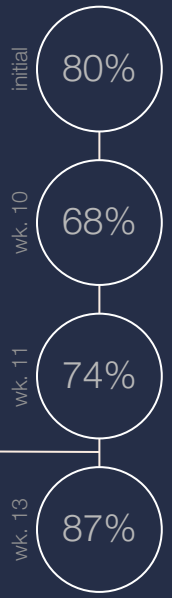
	Autoimmune	Benign	Skin Condition	Malignant	Infection	Unclassified
initial	2%	84%	61%	40%	71%	93%
wk. 10	x	75%	67%	64%	65%	76%
wk. 11	x	79%	74%	70%	70%	78%

model

Improvements

Gridsearch
- LR, transforms, optimizer, undersampling rates

Add data to underrepresented class - malignant



	Autoimmune	Benign	Skin Condition	Malignant	Infection	Unclassified
initial	2%	84%	61%	40%	71%	93%
wk. 10	x	75%	67%	64%	65%	76%
wk. 11	x	79%	74%	70%	70%	78%
wk. 13	x	89%	87%	86%	84%	89%

model

Improvements

Add Diverse Data

Combine image and metadata models into ensemble

		Autoimmune	Benign	Skin Condition	Malignant	Infection	Unclassified
initial	80%	2%	84%	61%	40%	71%	93%
wk. 10	68%	x	75%	67%	64%	65%	76%
wk. 11	74%	x	79%	74%	70%	70%	78%
wk. 13	87%	x	89%	87%	86%	84%	89%
final	85%	x	85%	80%	87%	81%	91%

evaluation

Performance

- (H) Notable improvement from baseline models
- (M) Ensemble Model Accuracy > 95%
- (M) Image Model Accuracy Accuracy > 90%
- (L) Deployed App Accuracy > 90%
- (H) Balanced classification across classes
- (H) High sensitivity to classifying the most dangerous categories wrong

Inclusion

- (M) Include diverse data to ensure model is usable for all skin types

Success Level
- high
- medium
- low

recommendation

Data

more out-of-sample test data

wider variety of image types across classes

Final Deployment

expand app to include metadata input

focus on out-of-sample failures in deployment

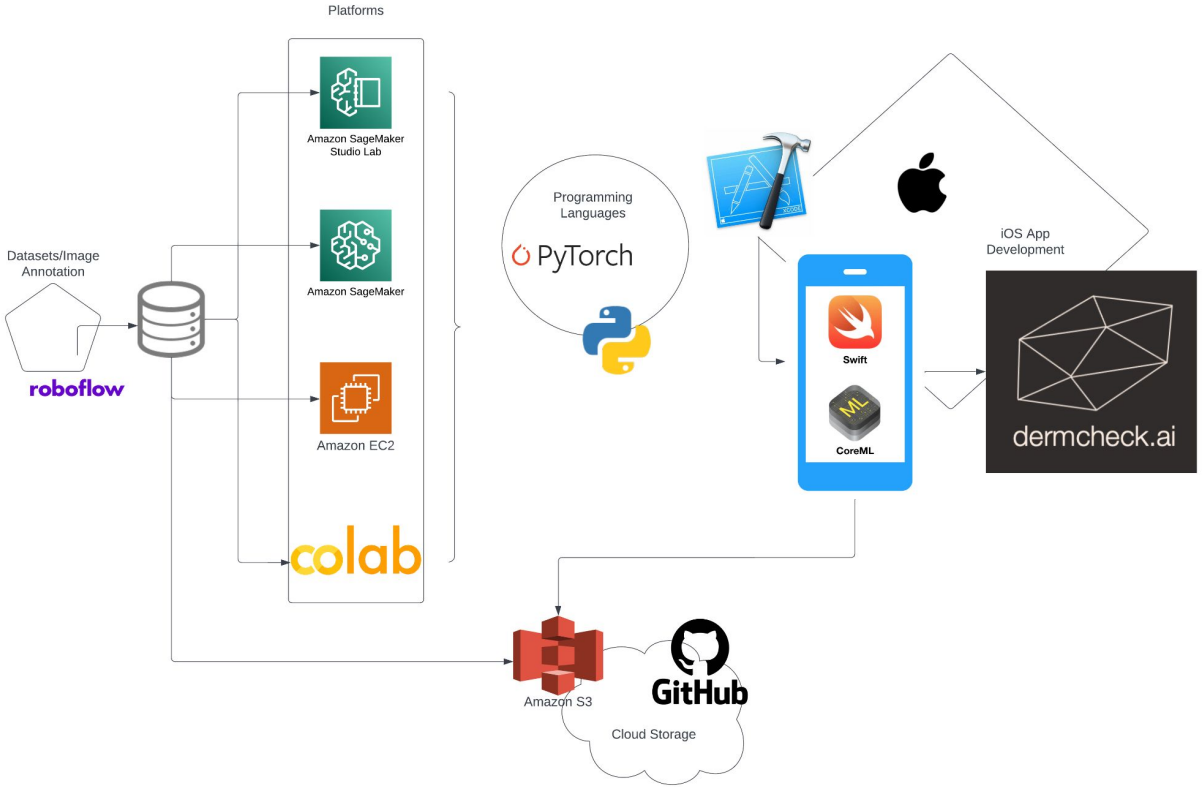
Inclusivity

increase the amount of skin types included in data

research state-of-the-art methods for adding more diverse images

dermcheck.ai

system architecture



evaluation

Performance

○ high success Notable improvement from baseline models	△ med. success Ensemble model accuracy > 95%	△ med. success Image model accuracy > 95%
□ low success Deployed app accuracy > 95%	○ high success Balanced classification across classes	○ high success High sensitivity to classifying the most dangerous categories wrong

Inclusion

△ med. success

Include diverse data to ensure model is usable for all skin types

Success Level
- ● high
- ▲ medium
- ■ low

dermcheck.ai

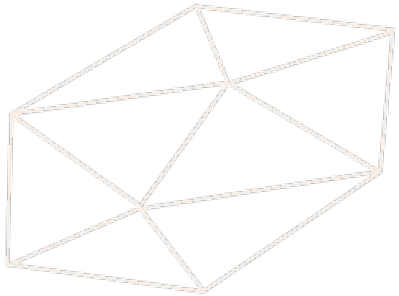
Final Presentation - Dry Run

Week 15

Ruby Han, George Jiang, Gerrit Lensink, Shivani Sharma

dermcheck.ai

dermcheck.ai



dermcheck.ai

Template