Team 7:

Do We Need Personalization and Big Data?

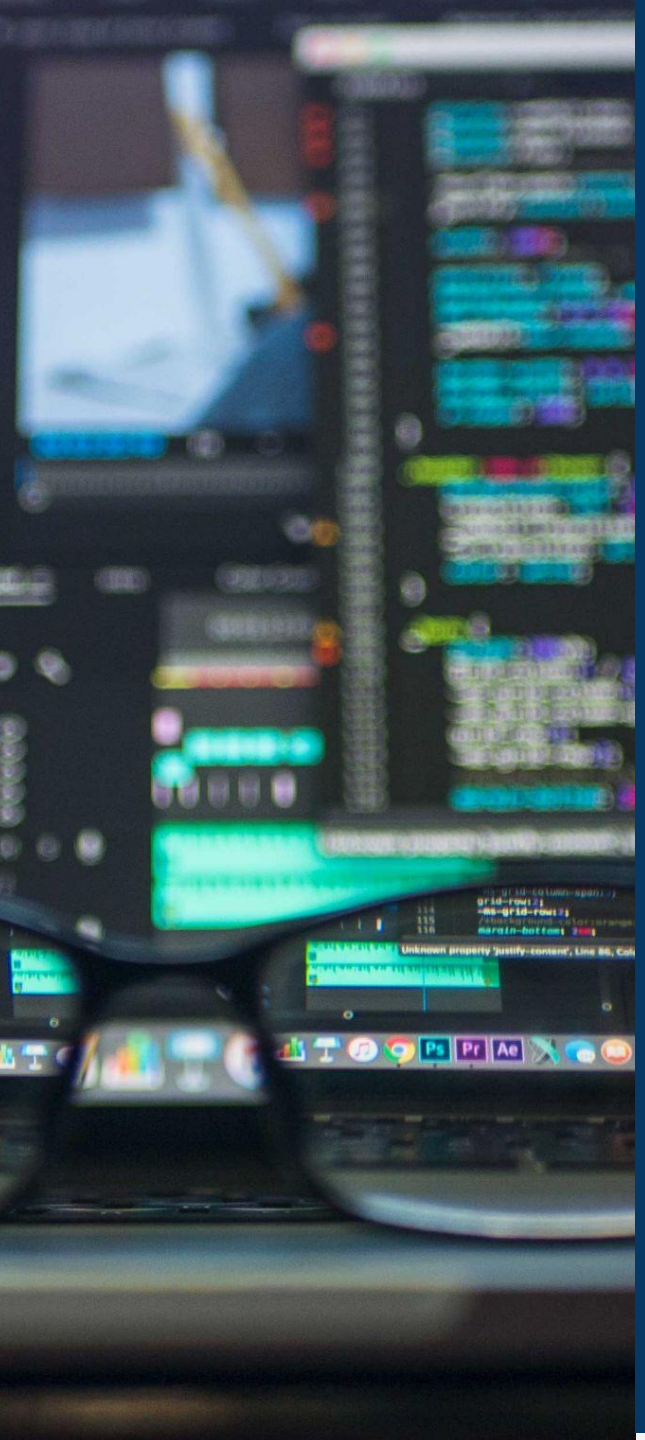
Yi Ge
Sadia Afreen
Francis Tuluri
Juan Luis Santana

Mentors:

Jack Gallifant and Po-Chih Kuo

Problem definition

- Can personalization with group attributes improve the performance at a group level?
- How does personalization affect subgroup performance compared to generic models?
- **Fair Use-** Personalised models should result in improved performance to predict no finding across all subgroups compared to the generic model



Dataset

- Emory CXR dataset
  - CXR Images
  - datathon_cxr_metadata.csv
  - datathon_cxr_findings.csv
- Demographic features: sex, age, race
- 70% training, 30% testing

Table 1

- 47.6% (White), 46.8% African American
- 45.4% Female, 54.6% Male

No Findings

		Missing	Overall	0.0	1.0	P-Value
n			7548	6319	1229	
Sex, n (%)	Female	0	3428 (45.4)	2815 (44.5)	613 (49.9)	0.001
	Male		4120 (54.6)	3504 (55.5)	616 (50.1)	
Race, n (%)	African American or Black	0	3531 (46.8)	2971 (47.0)	560 (45.6)	0.065
	Asian		137 (1.8)	120 (1.9)	17 (1.4)	
	Caucasian or White		3593 (47.6)	2979 (47.1)	614 (50.0)	
	Hispanic or Latino		55 (0.7)	47 (0.7)	8 (0.7)	
	Native Hawaiian or Other Pacific Islander		11 (0.1)	7 (0.1)	4 (0.3)	
	Unknown		221 (2.9)	195 (3.1)	26 (2.1)	
BMI_value, mean (SD)		0	28.7 (10.0)	28.6 (10.1)	29.4 (9.6)	0.006
Age, mean (SD)		0	56.8 (17.0)	58.0 (16.8)	50.6 (16.8)	<0.001

Models

- Association with demographics alone
- Densenet 121-> create embedding + probability of no finding
- Random forest with additional features
  - age
  - sex
  - race
- Evaluated subgroup performance on each combination
  - AUC
  - Personalization gain
  - Best/worst change in subgroup
  - Ratio gain/viol

Results

Features	Accuracy	F1 Scores
Age	0.72	0.80
Age, Race_White	0.73	0.80
Age, Race_White, Race_Black	0.74	0.80
Age, Race_White, Race_Black, Sex_Male	0.73	0.80
Age, Race_White, Race_Black, Sex_Male, BMI Value	0.86	0.87
Age, Race_White, Race_Black, Sex_Male, BMI Value, Insurance	0.87	0.88

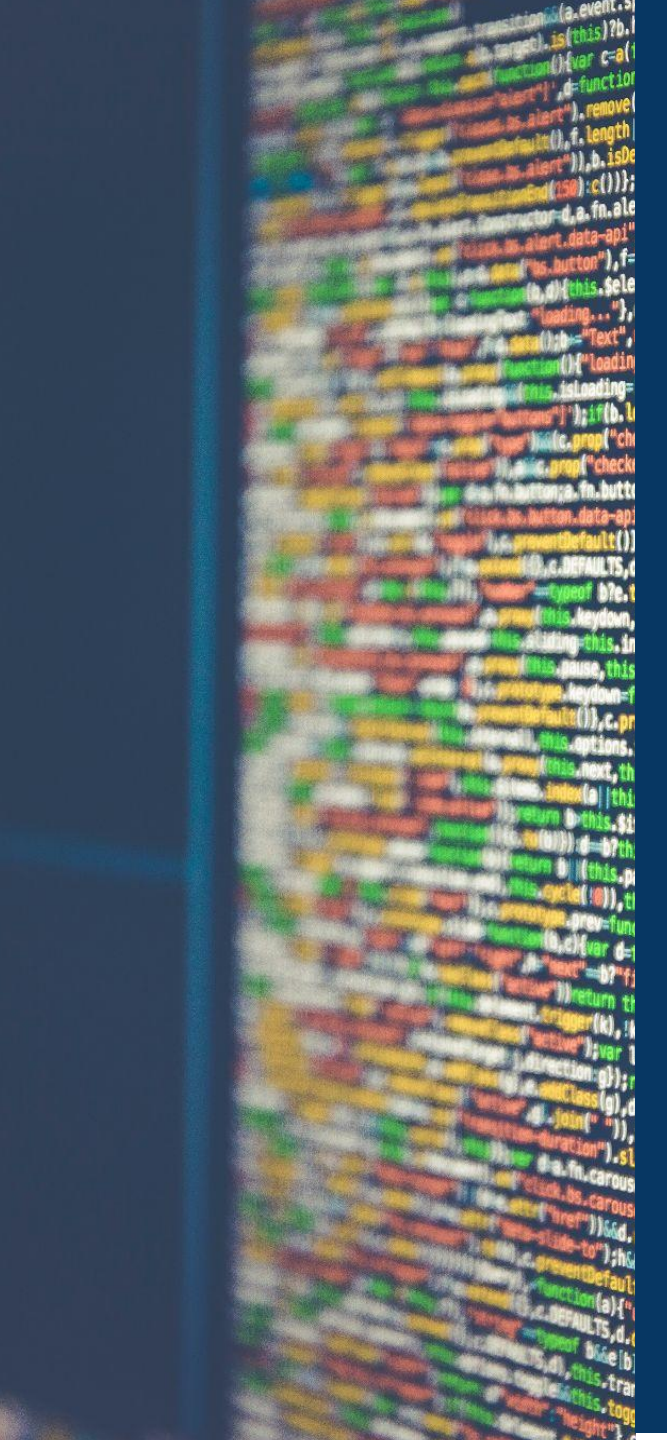
Results

- AUC of baseline model (NN) using only CXRs

Overall = 0.544	
Age	
-Young (<72)	0.537
-(Old) (>72)	0.585
Sex	
-Male	0.529
-Female	0.568
Race	
-(White)	0.550
-Non-(White)	0.537
-(Black)	0.527
-Non-(Black)	0.560

Results

	NN +Age	NN +Male	NN +White	NN +Black	NN +Age +Male	NN +Age +White	NN +Age +White	NN +White +Male
AUC	0.537	0.530	0.536	0.528	0.531	0.558	0.557	0.545
AUC Gain	-1.3%	-2.6%	-1.5%	-2.9%	-2.4%	+2.6%	+2.4%	+0.2%
Best Gain (Best Group)	+3.4% (Black)	+2.9 (White)	+2.1 (Black)	+1.5 (Black)	+2.2 (White)	**+9.5 (Black)**	+3.6 (Black)	+2.9 (White)
Worst Gain (Worst Group)	-12.5% (Old)	-7.5% (Non-White)	-8.7% (Old)	-7.9% (Old)	**-20.5% (Old)**	-6.0% (Old)	-3.8% (Old)	-3.9% (Old)
Ratio gain/viol	3/5	2/6	1/7	1/7	1/7	5/3	6/2	4/4



Discussion

- Fair use calculations can be calculated
- Personalization may result in overall performance boost
- Personalization does not always result in benefits to subgroups

Future work

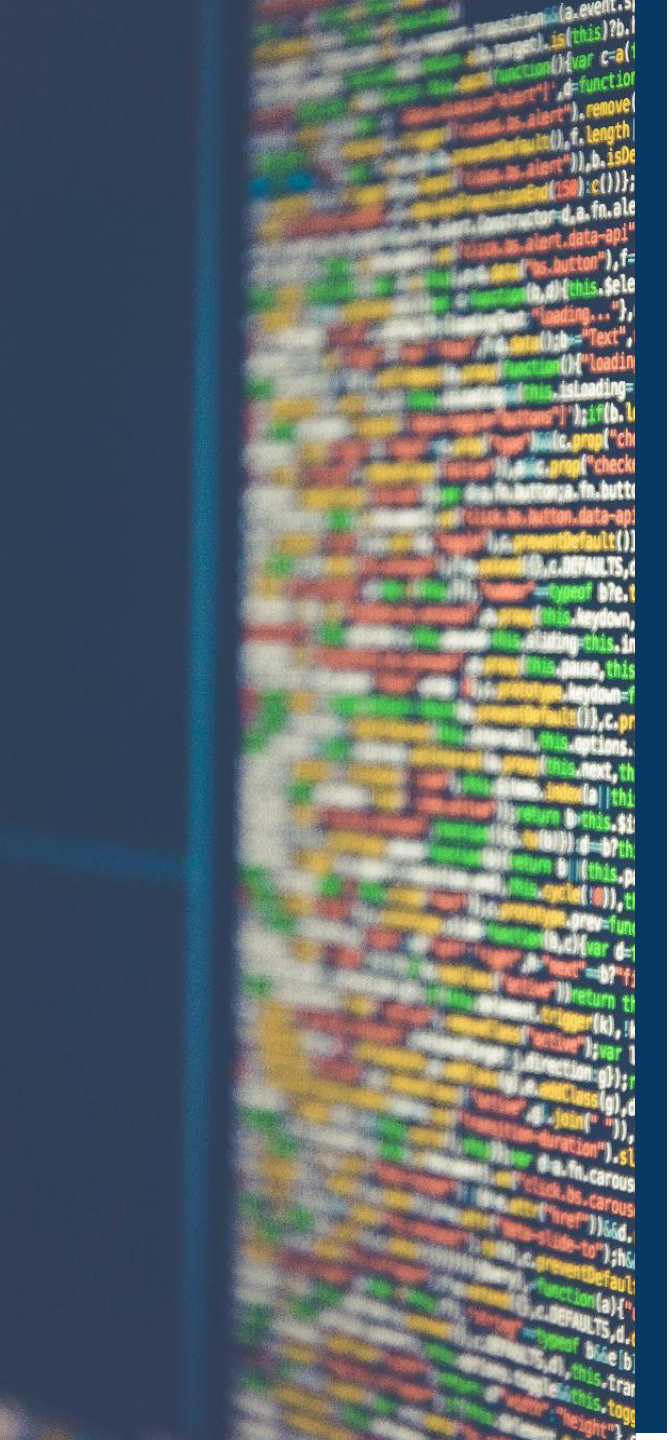
- Fine tuning
- Other diagnoses
- Other models
- External validation

Summary- Do We Need Personalization and Big Data?

- **Fair Use-** Personalised models should result in improved performance to predict no finding across all subgroups compared to the generic model

- Evaluation of Image embedding + Random forest with additional features

	NN +Age	NN +Male	NN +White	NN +Black	NN +Age +Male	NN +Age +White	NN +Age +White	NN +White +Male
AUC	0.537	0.530	0.536	0.528	0.531	0.558	0.557	0.545
AUC Gain	-1.3%	-2.6%	-1.5%	-2.9%	-2.4%	+2.6%	+2.4%	+0.2%
Best Gain (Best Group)	+3.4% (Black)	+2.9 (White)	+2.1 (Black)	+1.5 (Black)	+2.2 (White)	**+9.5 (Black)**	+3.6 (Black)	+2.9 (White)
Worst Gain (Worst Group)	-12.5% (Old)	-7.5% (Non-White)	-8.7% (Old)	-7.9% (Old)	**-20.5% (Old)**	-6.0% (Old)	-3.8% (Old)	-3.9% (Old)
Ratio gain/viol	3/5	2/6	1/7	1/7	1/7	5/3	6/2	4/4



Results: Random Forest Model

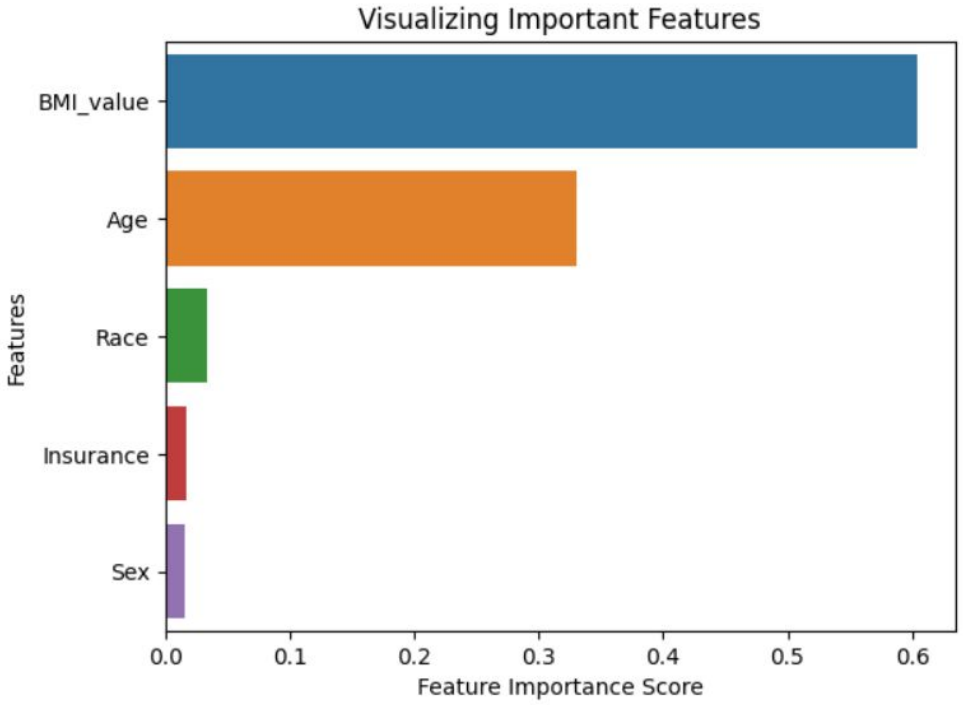
- Features and target

	Sex	Race	BMI_value	Age	Insurance
0	0	1	26.7	80	1
1	0	2	23.9	55	1
2	0	2	23.9	55	1
3	1	1	20.1	77	1
4	1	1	20.1	77	1
5	1	1	31.3	78	0
6	1	1	27.2	61	1
7	1	2	44.8	59	1
8	1	2	44.8	59	1
9	0	2	24.6	47	0

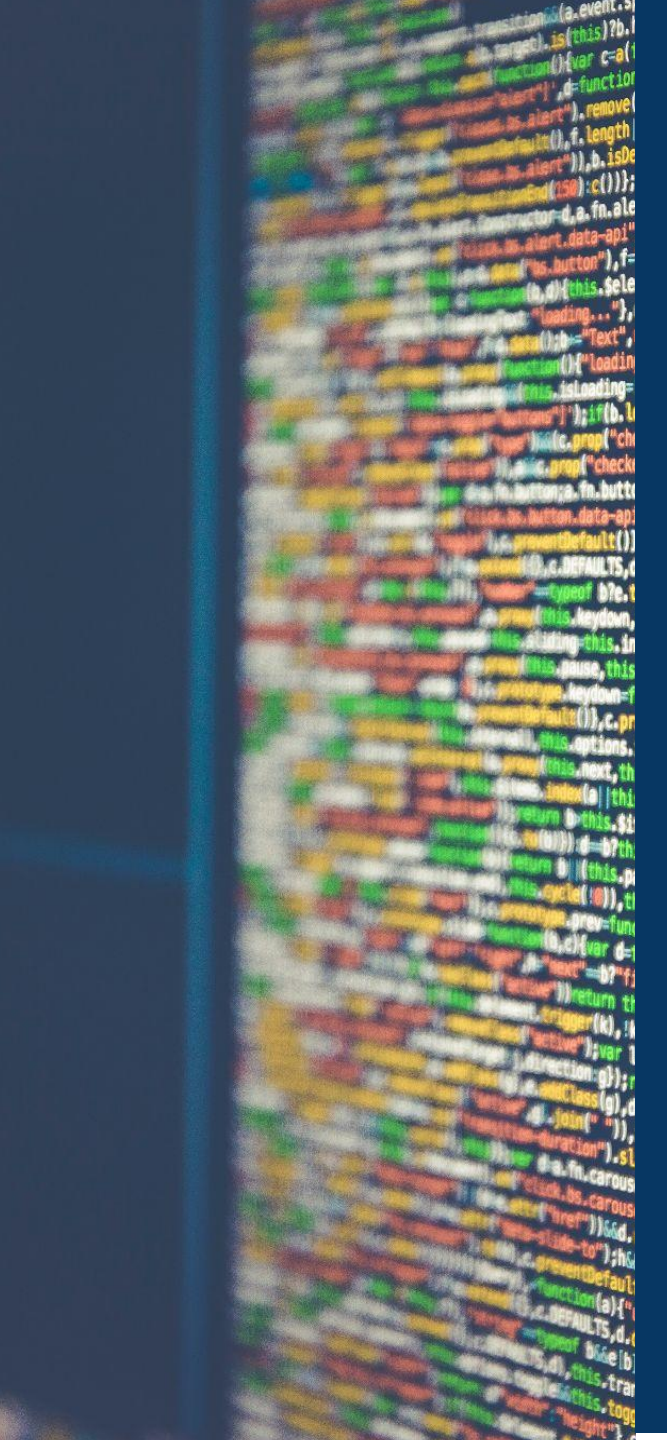
	Pneumonia
0	0.0
1	0.0
2	0.0
3	0.0
4	0.0
5	0.0
6	0.0
7	0.0
8	0.0
9	0.0

Results: Random Forest Model

- 100 trees; Feature importance




```
BMI_value     0.604663
Age           0.329989
Race          0.033569
Insurance     0.016388
Sex           0.015391
dtype: float64
```



Results: Random Forest Model

- 100 trees; Metrics

	precision	recall	f1-score	support
-1.0	0.52	0.34	0.41	196
0.0	0.92	0.96	0.94	2505
1.0	0.57	0.43	0.49	147
accuracy			0.89	2848
macro avg	0.67	0.58	0.62	2848
weighted avg	0.88	0.89	0.88	2848