

Predicting Discontinuation of Docetaxel Treatment for Metastatic Castration-Resistant Prostate Cancer (mCRPC) with Hill-Climbing and Random Forest

Daniel Kristiyanto, Kevin Anderson, Seyed Sina Khankhajeh, Kaiyuan Shi, Seth West, Ling Hong Hung, Azu Lee, Qi Wei, Migao Wu, Yunhong Yin, Ka Yee Yeung.

Correspondence Email: kayee@uw.edu

BACKGROUND

- #1 CANCER IN MEN & AND ONE OF LEADING CAUSES OF CANCER MORTALITY
- TREATMENT: RADIO THERAPY & ANDROGEN DEPRIVATION THERAPY (ADT)
- PATIENTS DEVELOPED RESISTANCE TO ADT: METASTATIC CASTRATE-RESISTANT PROSTATE CANCER (MCRPC)



INPUT

131 CLINICAL VARIABLES

MEMORIAL SLOAN KETTERING (MSK)
476 PATIENTS

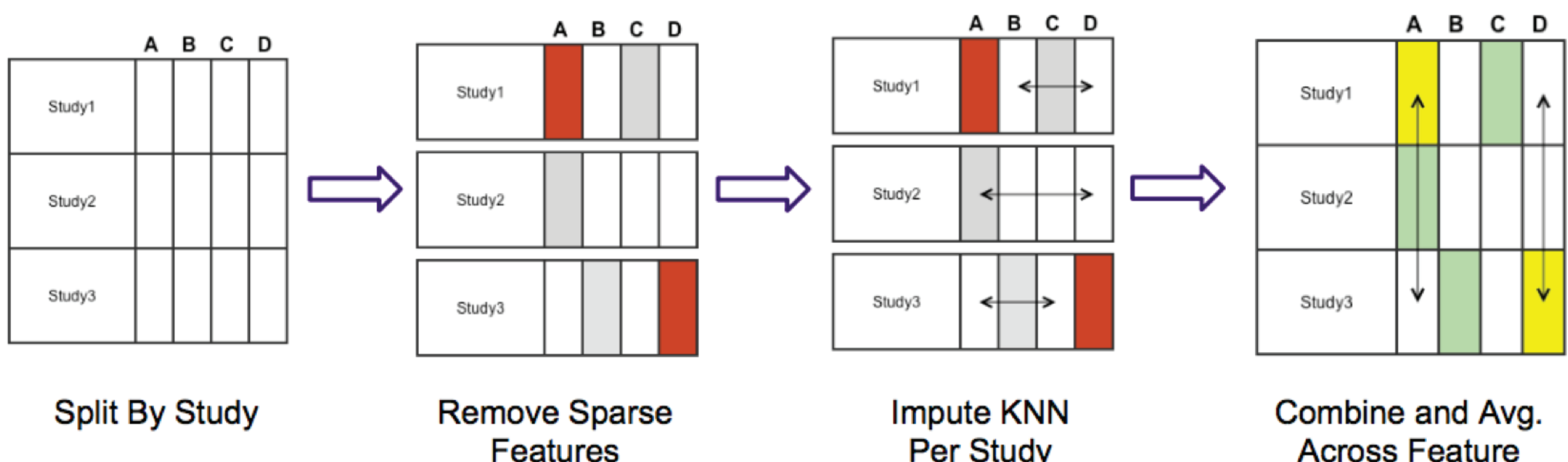
CELGENE
526 PATIENTS

SANOFI
598 PATIENTS

ASTRAZENECA (AZ)
470 PATIENTS

GOAL: PREDICT DISCONTINUATION OF TREATMENT FOR AZ DUE TO ADVERSE EVENT WITHIN 3 MONTHS

DATA & METHODS



DATA CLEANSING DONE SEPARATELY WITHIN EACH CLINICAL TRIAL, AND MERGED BACK TOGETHER.

1) REMOVAL OF CLINICAL VARIABLES

2) IMPUTATION

3) MULTILABEL INTO BINARY VARIABLES

FEATURES SELECTION:



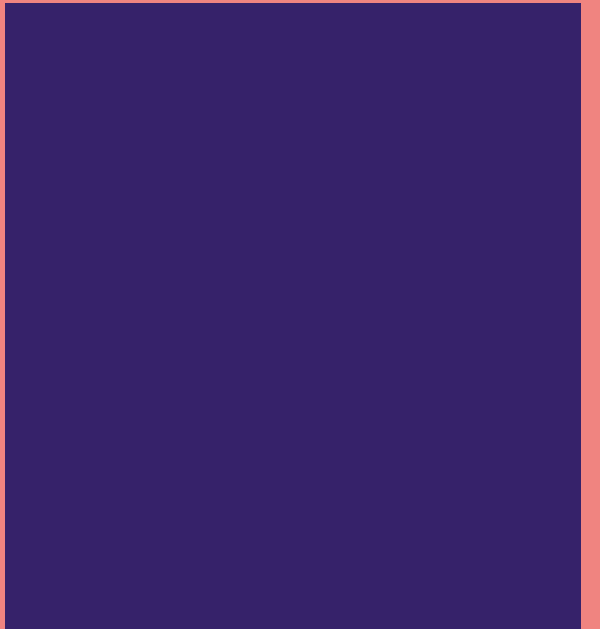
UNIVARIATE: AUC 1.2



HILL CLIMBING METHOD

DATA BALANCING:

1292 NEGATIVE SAMPLES



197 POSITIVE SAMPLES



RANDOM SAMPLE

60% NEGATIVE

40% POSITIVE

CLASSIFICATION

CELGENE

SANOFI

VALIDATION SET

MSK

TRAINING SET



Random forest using Sanofi as the holdout

RESULTS

ACROSS THE 470 IN ASTRAZENECA CLINICAL TRIAL, 8 PATIENTS ARE PREDICTED TO DISCONTINUE THE TREATMENT WITHIN 3 MONTHS.

ASSESSMENT

Final Scoring (AZ)

Training Set (MSK + CELGENE + SANOFI)

0.129

AUC

0.47

ACCURACY

0.89

PRECISION

0.8

10 CROSS-FOLD VALIDATION
(MSK + CELGENE + SANOFI)

F1

0.33

RECALL

0.21

0.131

AUC

0.165

ACCURACY

0.9

PRECISION

0.21

ADDITIONAL ASSESSMENT
(MSK AND CELGENE AS THE TRAINING SET)

F1

0.092

RECALL

0.06