Machine Learning Approaches in Systems Biology Modeling

- Learning approaches in data processing – Data Mining

- Decision Making systems for Pathways modeling
  - or How ML synergies with deterministic modeling

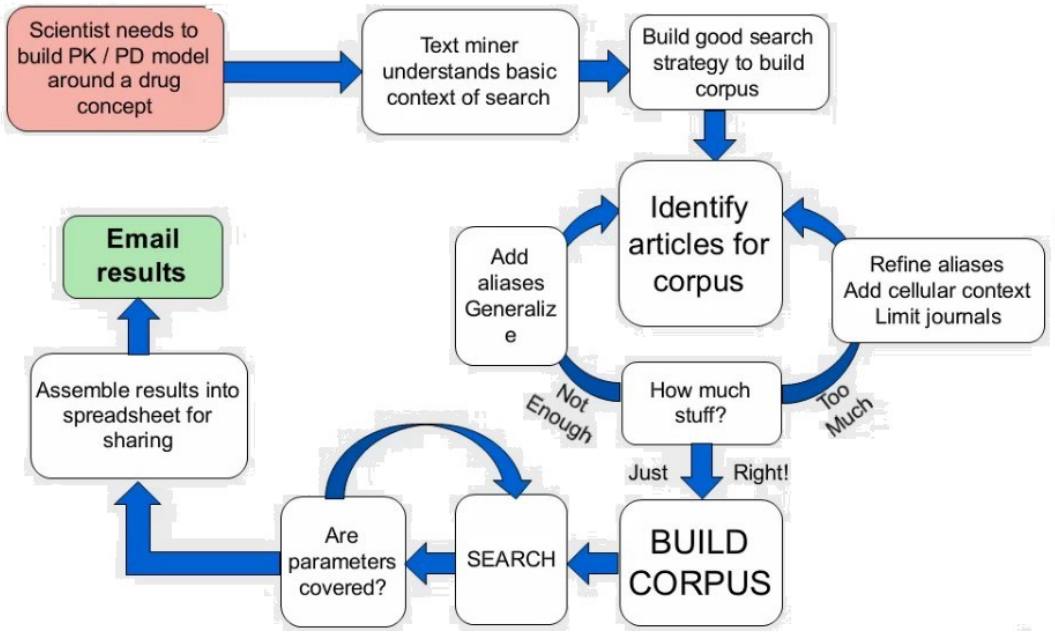
- Simplified Classification Systems

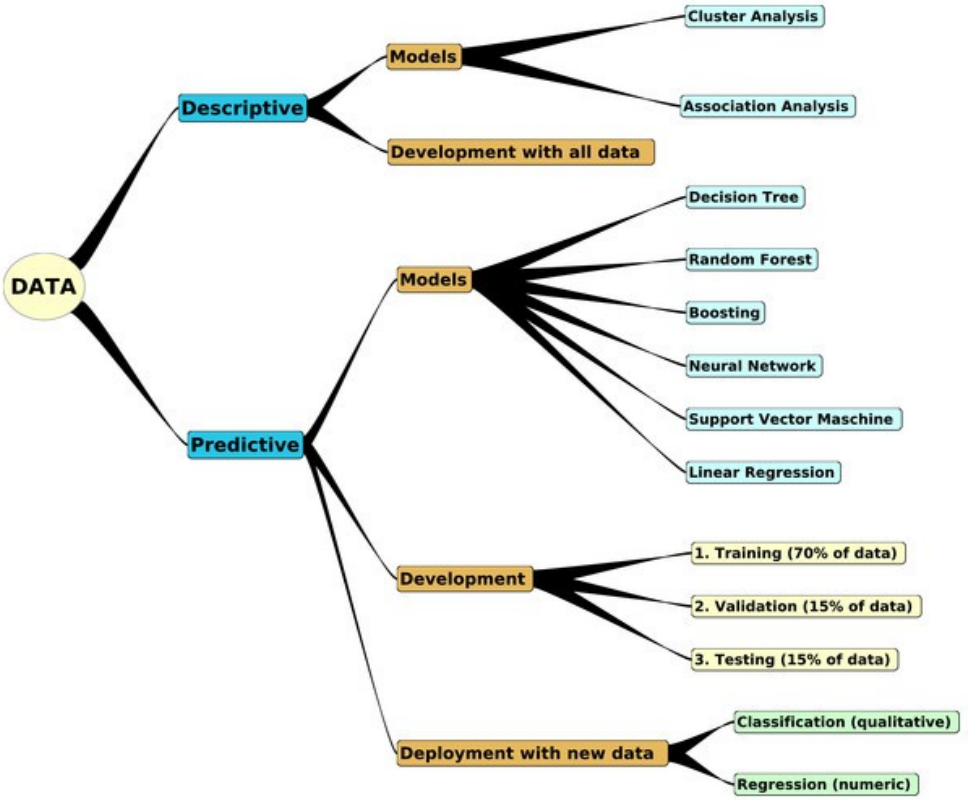
- Limitations of pattern-based decision making
  - Genomic data-loss in pattern-building, loss of uniqueness and personalization

- New possibilities of Stochastic-Deterministic synergism (SDS) – cancer susceptibility prediction systems

Learning approaches in data processing – Data Mining

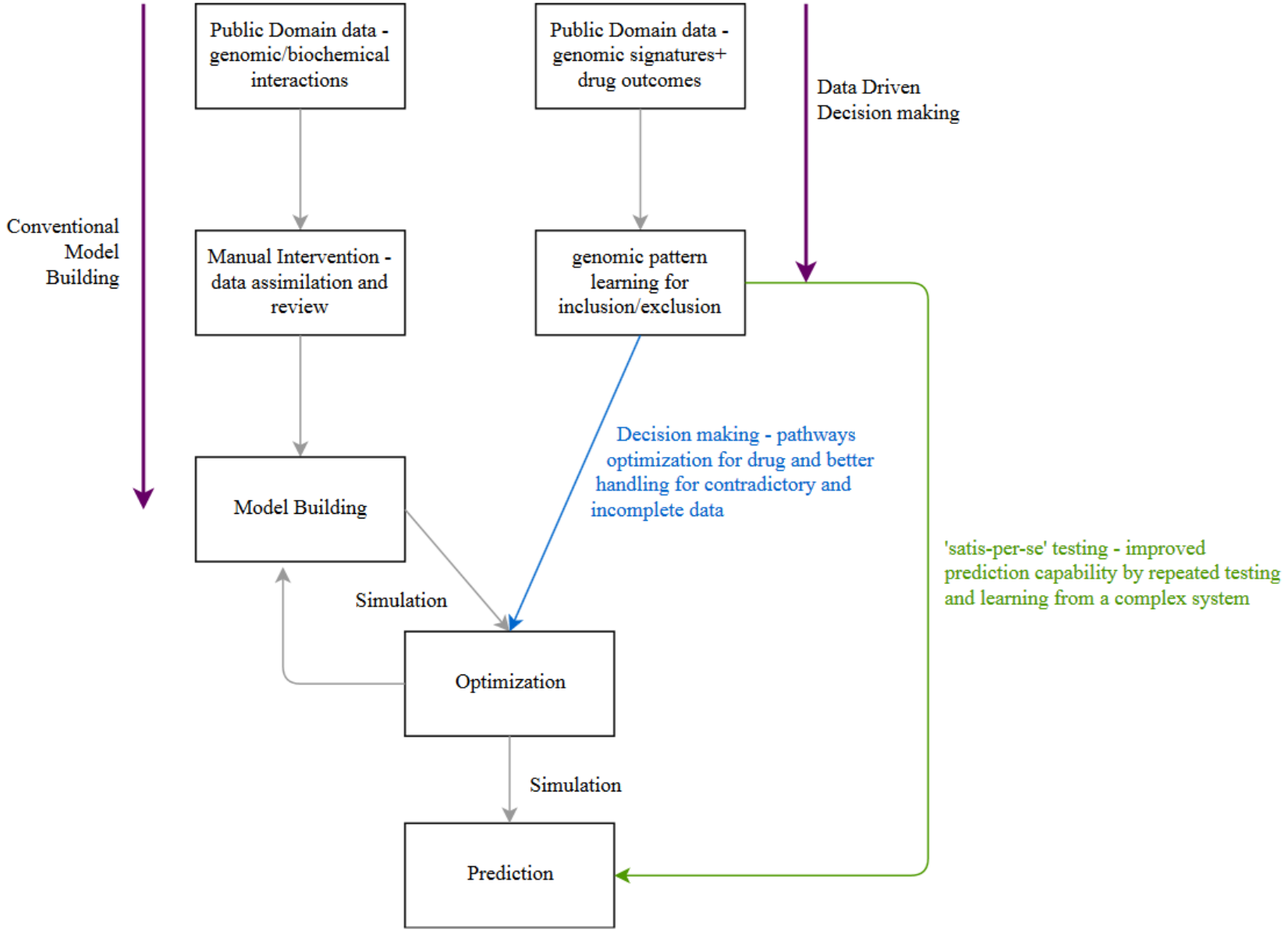
Source : peerj.com



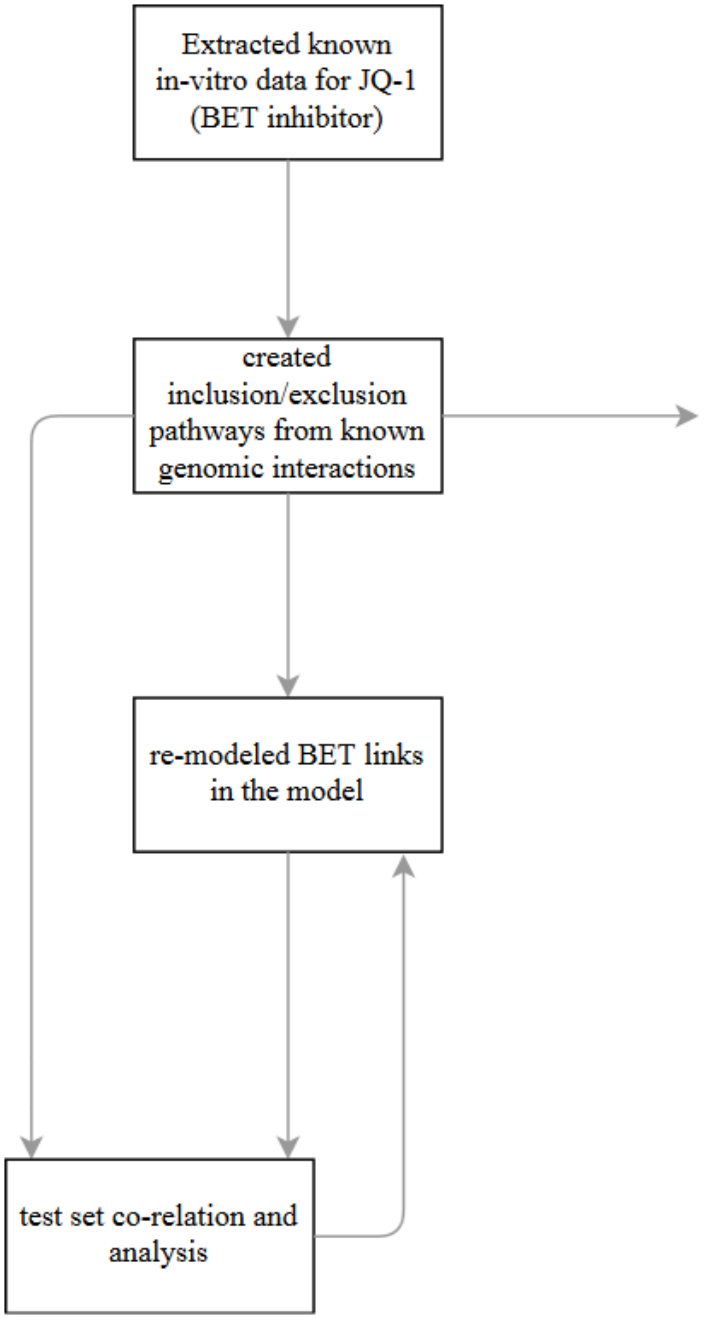


Decision Making systems for Pathways modeling

- or How ML synergies with deterministic modeling

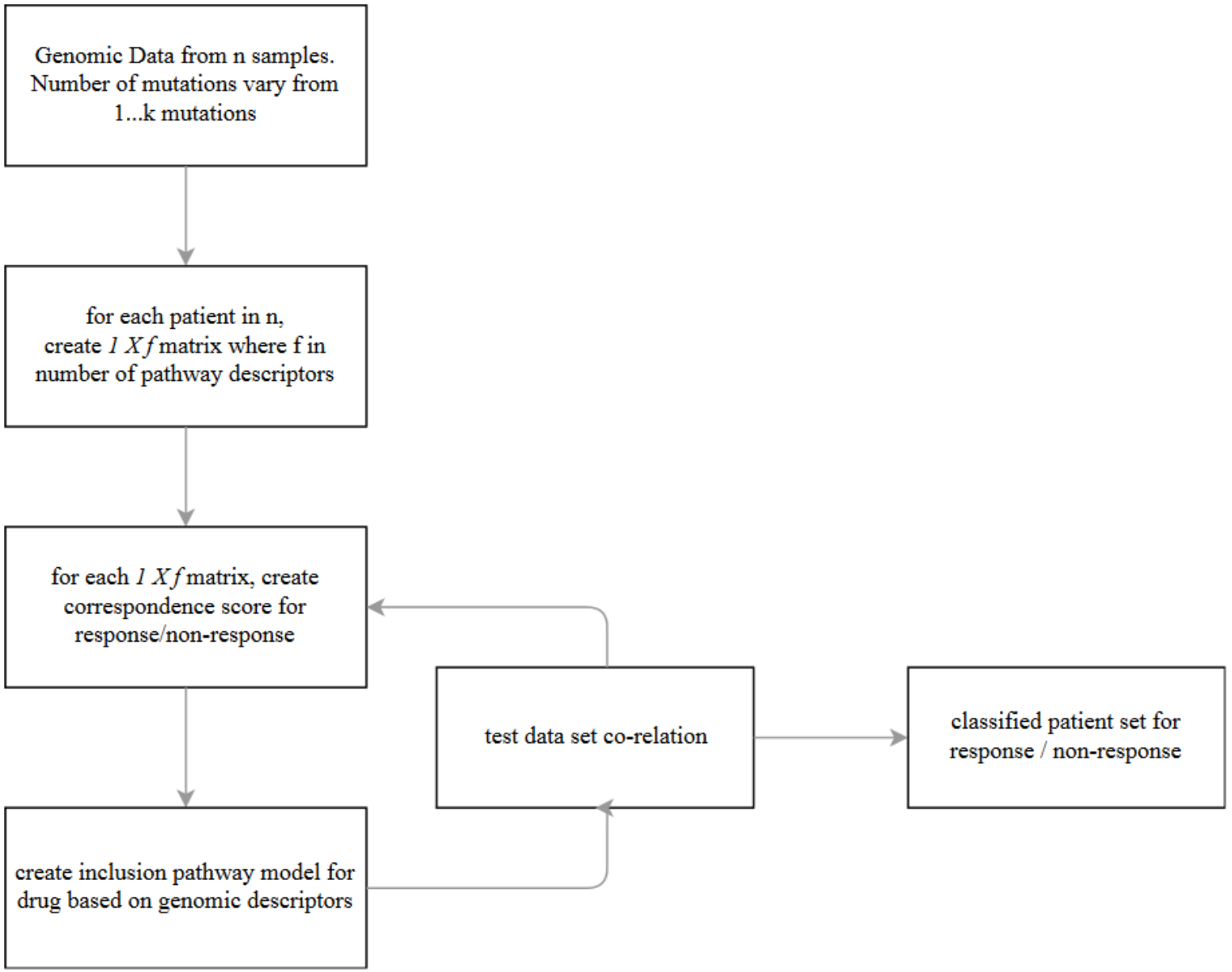


Case 1 : JQ-1 Drug Modeling

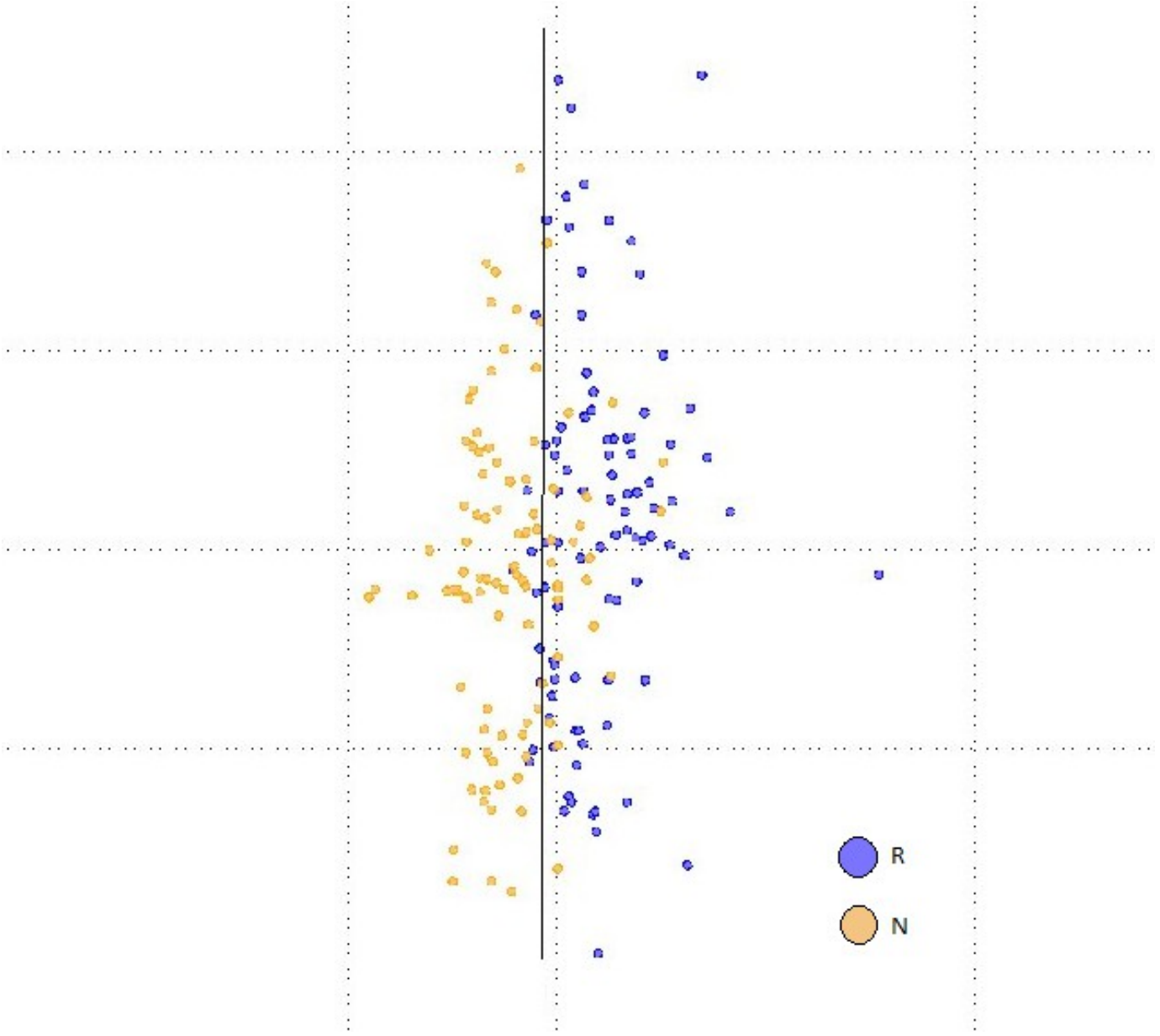


NEGATIVE FACTORS	
APOPT	-4.6948170597
P53P	-2.8447406455
MAPKSIG	-2.0924796688
BVEGFSIG	-1.0478771424
HRP	-0.982384821
NFKBP	-0.4929414283
NOTCHSIG	-0.3516315522
MTORP	-0.2991624558
TNFSIG	-0.243184438
PLATR	-0.2366118856
RASSIG	-0.2218236427
TLRSIG	-0.206913686
ESRSIG	-0.1863318599
NHEJP	-0.1424053015
PLDSIG	-0.1314510476
SPLICE	-0.1202112918
HIPPOSIG	-0.1081901626
SPILOM	-0.1068544816
CHOLM	-0.1009774851
CAMPSIG	-0.0919616382
HYPOXSIG	-0.0801408612
FOXOSIG	-0.0793394526
ARACM	-0.0615622528
PROTES	-0.0544166341
MMRP	-0.0540950813
BCRSIG	-0.0414061116
PEROX	-0.0321704894
CPKGSIG	-0.0293153585
FRU-MANM	-0.0260580964
PYRUVM	-0.0260580964
FOLDP	-0.0247125844
TRPM	-0.0244294654
PYRM	-0.0234522868
GLYCOL	-0.0193022936
RIGLSIG	-0.00882923
PURM	-0.0065377208
TCRSIG	-0.00573685
FATSYN	-0.0038742049
PPP	-0.0028697814
POSITIVE FACTORS	
H3K27ME3P	-0.0017271833
VB6M	-0.0010935
ARG-PROM	-0.0010206
ERBBSIG	0.0494201977
HEDGESIG	0.0757331138
GLUGOSIG	0.0778262996
INSPATH	0.32341887 11
EPIGEN	0.7422463092
CELCYS	9.2505974105

Simplified Classification Systems



Case 2 : Methotrexate Response Classification



Limitations of pattern-based decision making

- Genomic data-loss in pattern-building, loss of uniqueness and personalization

Genomic Data-Loss

Current pattern-recognition approach relies on converting the n dimensional genomic mutations data, where *n* is the number of mutations, into a 1 X f matrix, where f is a fixed number, and represents the pathways descriptors. Hence, this is a *generalization* of the actual mutation, which represents more the general class of patient population rather than the individual patient

Loss of personalization

This results in a lack of uniqueness in the data and hence a loss of personalization . This can also lead to loss of unique mutation factors that are not accounted for in known literature or in pathway descriptors. This can be achieved only through recurrent occurrence and/or scientific literature.

However, with deterministic modeling, this can be easily established and proven via simulations. *In other words, data-mining / pattern recognition approaches can identify patterns and establish hypotheses, but cannot create new data. A combination of deterministic and stochastic approaches can accomplish this at a faster rate and higher frequency.*

Limitations of pattern-based decision making

- Genomic data-loss in pattern-building, loss of uniqueness and personalization

Reduced Focus

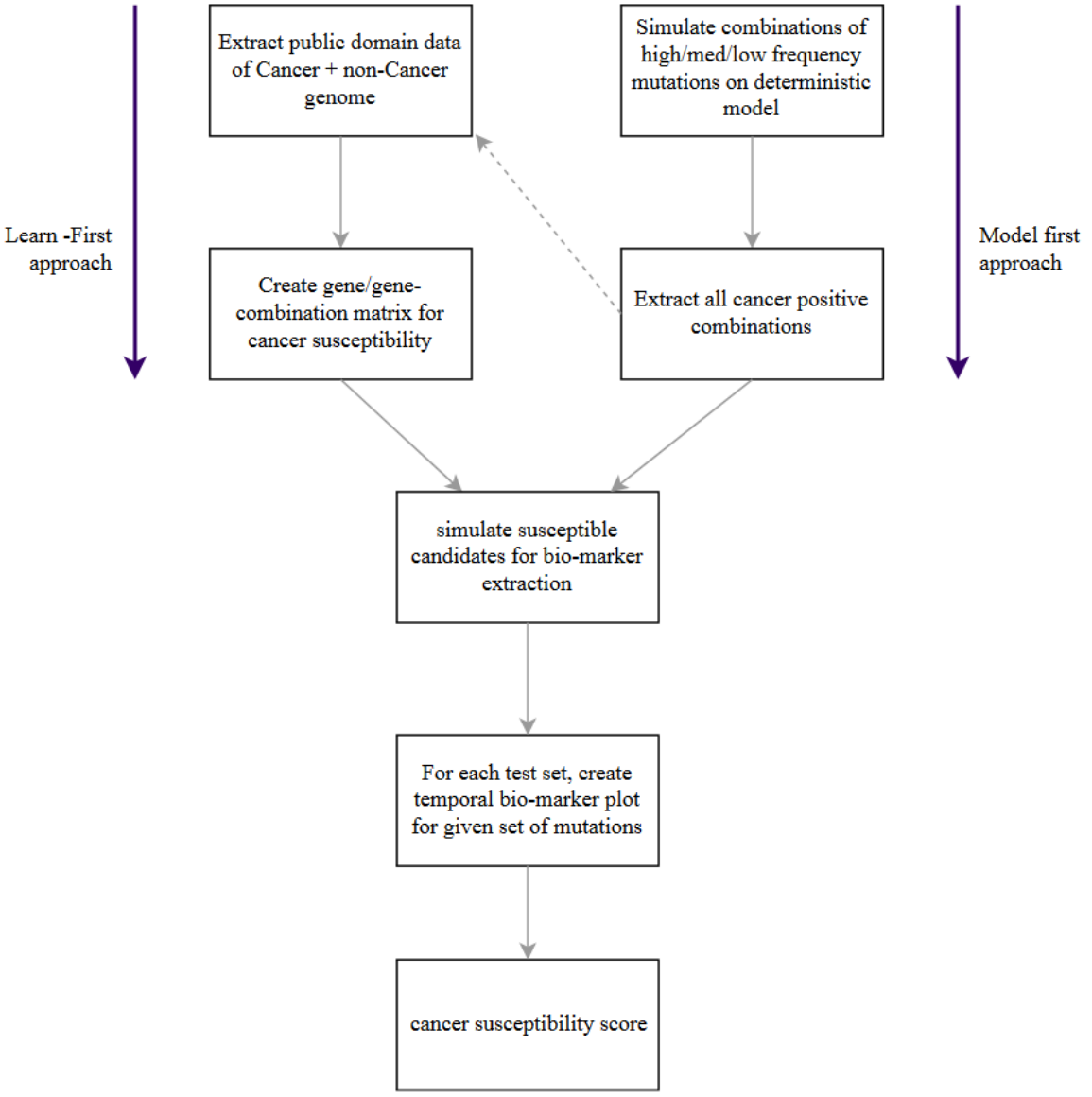
Based on genomic descriptors of a patient, we can construct a generalized pathway description of the patient cancer-biome. Since all the drug inclusion and exclusion are defined based on the same set of descriptors, individual drugs cannot be identified as best-fit for the patient, rather a normal curve for the general class of drugs that can be the probable best fit can be generated.

This leads to reduced focus of drugs which can be identified and proved by deterministic simulations

Hence pattern-recognition here can serve as a filter for selecting best class of drugs based on genomic generalizations but the best drug/combination of drugs can be selected based on simulation data

New possibilities of Stochastic-Deterministic Synergism (SDS)

– cancer susceptibility prediction systems



New possibilities of Stochastic-Deterministic Synergism (SDS)

Virtual clinical trials