Contextual Bandits Decision Tree:

	Opens	Clicks	Mean
A	4000	140	0.035
B	3000	150	0.05

X	Option	Opens	Clicks	Mean
0	A	2500	50	0.02
	B	1000	100	0.10

X	Option	Opens	Clicks	Mean
1	A	1500	90	0.060
	B	2000	50	0.025

Decision Node: ●

Strategy	Mean (μ)	Incremental Clicks
Action*	0.05	60
Action*|X	0.08	210

Value Selection:
The expected reward (μ) when using variable X_i to select the best action is the weighted sum of expected rewards of the best actions for each of its possible values
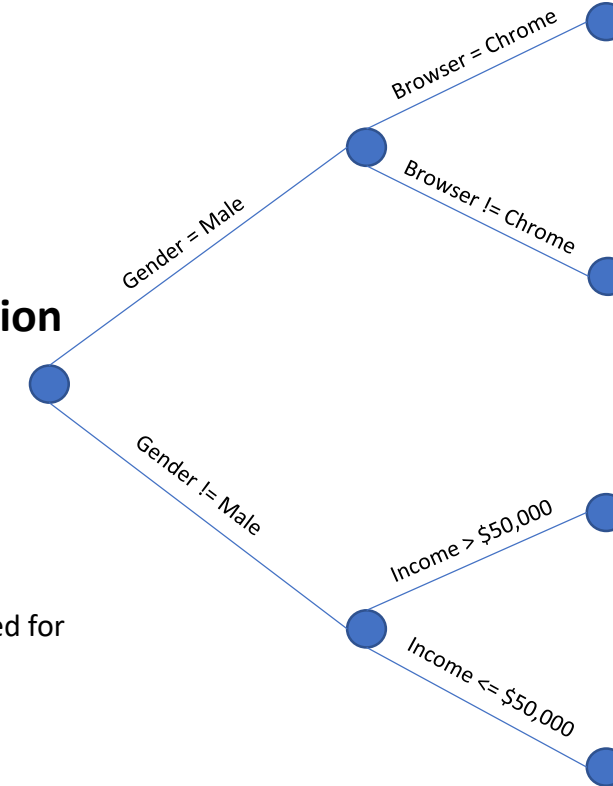
For continuous variables, a binary split is created to produce two segments of the data: > or <= value

For categorical variables, a one-vs-all strategy is used to create the two segments: = or != value

Variable Selection:
The optimal variable to be used for selecting the best action is:
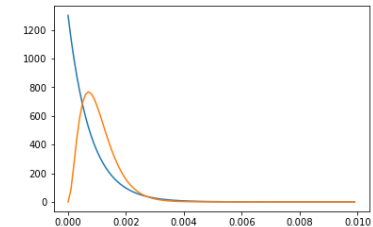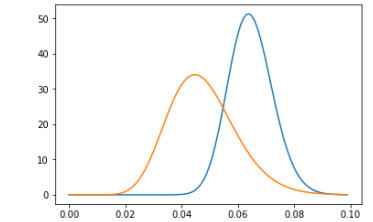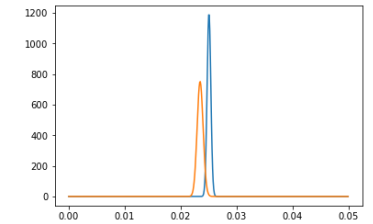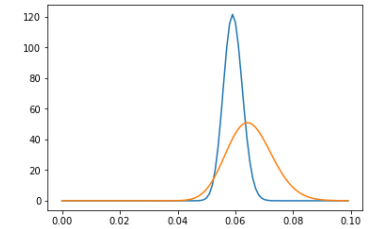$X_i * = \text{argmax}(\mu | X_i)$

Decision Tree:

Gender = Male
Browser = Chrome
Browser != Chrome
Gender != Male
Income > $50,000
Income <= $50,000

Thompson Sampling:
Each option is modeled by a Beta Distribution with:
b = #opens
a = #clicks

Original idea came from:
Feraud, R. , Allesarido, R. , Urvoy, T. & Clerot, F.
Random Forest for the Contextual Bandit Problem