Problem Statement

Predict Consumer Purchases

The data set attached to this task contains the purchase histories (i.e., shopping baskets) of

2,000 consumers over 49 weeks across 5 categories ( *train.csv* ). In simulating the basket data

we assumed that consumers only buy one unit of a product in a given week. The data set also

contains the price consumers paid for one unit of product j in week t and a boolean variable

that indicates whether the purchased product was advertised (1) or not (0). We also provide

the week 50 promotion schedule (discounts and advertising) for all products

( *promotion\_schedule.csv* ).

*Your task:*

Use the data to build a ML model for consumer purchases. With the trained model, predict

week 50 purchases for all 80,000 possible consumer-product combinations (40 products x

2,000 consumers) in the data. Feel free to use any non-parametric or “black box” model you

consider appropriate (i.e., you don’t have to follow the formalisation from part 1). Please

provide your predictions as a *.csv* file (cf. *prediction\_example.csv* ) that contains the columns i

(consumer), j (product), and prediction. We will benchmark your predictions against observed

purchases using the AUC metric.

i= Consumer id

j=Product category

t= Week(0-49) Predict for 50th

price=price of the product

advertised= (1=yes,0= no)