



Predictive Models Using Text and Numeric Data

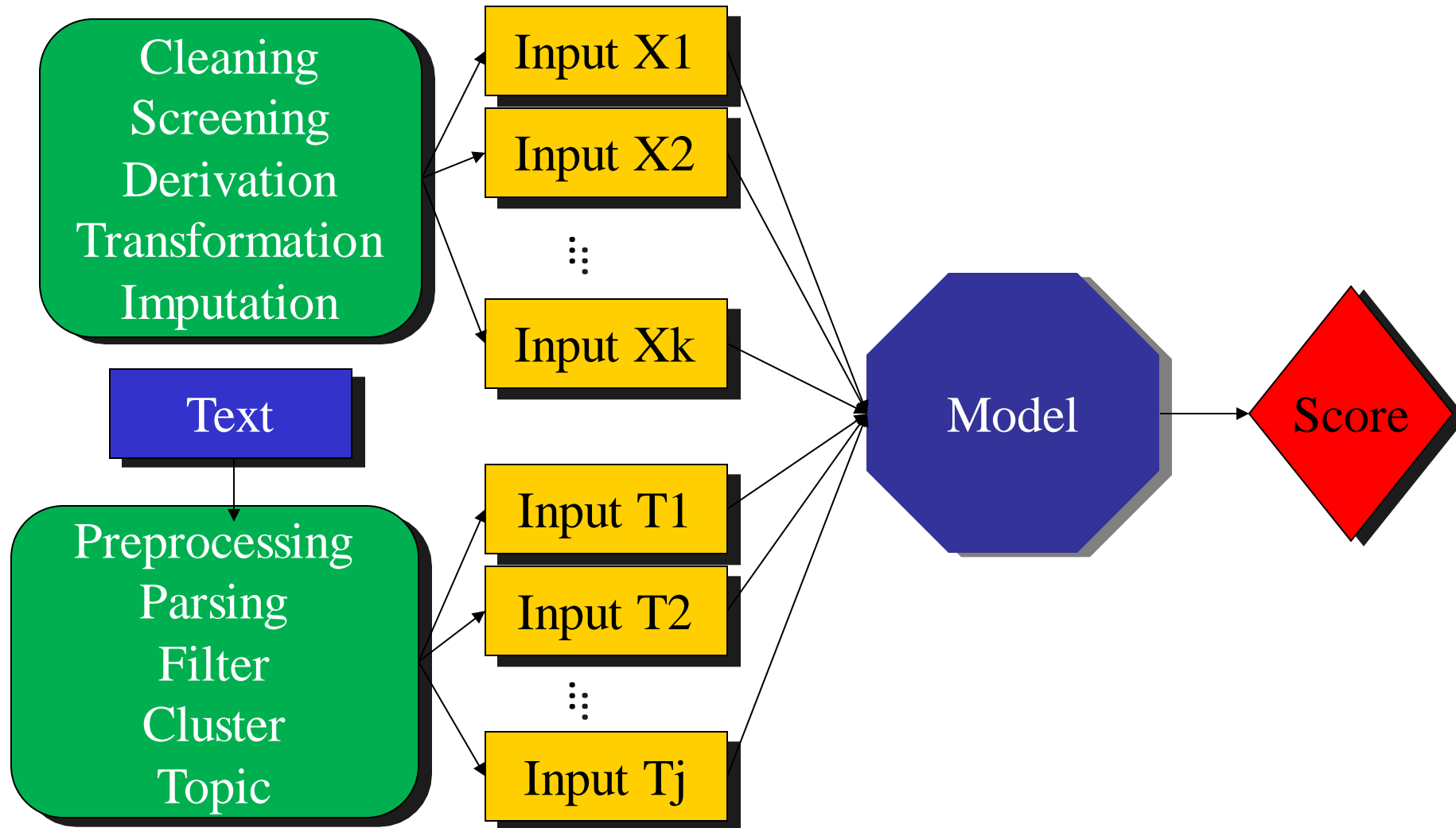
Dr. Goutam Chakraborty



Using Both Text and Numeric Data in Predictive Models

- Traditionally, predictive models are built with just numeric data. These models can often be improved by bringing in appropriate textual data.
- Examples:
 - Using customer survey comments to improve prediction of numeric variable in the survey such as customer satisfaction
 - Using call-center transcripts to improve prediction of churn models built on transaction data only

Predictive Modeling with SAS Text Miner





Using Text Data in Predictive Models

- Loyalty card holder customer (of a fuel-store) comments (call-center records) are used as text input.
- The model is built first with numeric data only to predict the target variable, a binary flag that reflects a type of future behavior (actual description masked for confidentiality).
- The model built with numeric plus text clusters (of comments) substantially outperformed the model built with numeric data only.



Procedure

- Follow handout titled “Demo with Call Center Data”