

Machine Learning Methods Applied to a Prescription Drug Marketing Problem

Data Mining Technologies, Inc.
mgilman@data-mine.com

Overview:

- Objective: Apply machine learning algorithms to Pharma Clients' historical Drug Prescriber set in an attempt to determine a Prescriber segmentation methodology that, when piloted in-market will demonstrably and materially grow annual Drug scripts written in-market
- Leverage actual observed Prescriber behaviors, and as such, enhance Pharma Clients' "4I's" and top 10% sales volume segmentation and targeting methodologies.
- Conduct low cost, live, in-market tests to validate the model and accelerate speed to positive market action

This methodology will assess, quantify, predict, and validate Prescriber propensities to write more Drug prescriptions in response to the targeted application of key promotional stimuli applied to the most responsive prescriber targets

Pharma Clients: Context

- **Current Physician Targeting Methodology**
 - Targets identified via volumetric analysis and statistical modeling
 - Sales Force: Makes Ad Hoc changes based upon sales representative experience and input
- **Challenge**
 - Extremely broad and shallow data set limits the effectiveness of traditional statistical modeling
- **Machine Learning Solution**
 - Apply advanced machine learning to data sets to augment traditional statistical modeling

Typical Model Building Process

- Define the problem ->
- Generate a flat file (i.e. training file)->
- Build classification model ->
- Validate the model ->
- Generate the model (rules set)-> Apply results by providing a personalized recommendation for each physician with accompanying auditable rule->
- The collective set of rules (i.e. the model) suggest the most effective marketing stimuli by physician to achieve this goal and also estimate the probability that the rule will work->
- Apply the rules to change marketing tactics for each physician, and measure success rates to validate in-market performance

Machine Learning Approach

- ✓ Supervised, Automated Learning Using AI and Machine LEARNING to Classify Physician Prescribing *Behavior*
 - No a priori assumptions (e.g. linearity, non-linearity, etc.) in relationships between Target and Input variables
 - Not beholden to typical complete-case analysis requirements
 - ✓ Automatic handling of missing values, multiple data input types (e.g. categorical, continuous, nominal), outliers
 - ✓ Multiple methods for missing value imputation
 - Both Automated and User-Defined variable binning
- ✓ Produces easy to understand and interpret business Rules w/Confidence Factors
- ✓ Outperforms traditional statistical methods against rare incident targets, sparse input values, & semi-structured data

Data Inputs

| Data extraction for IMS NPI number for physicians

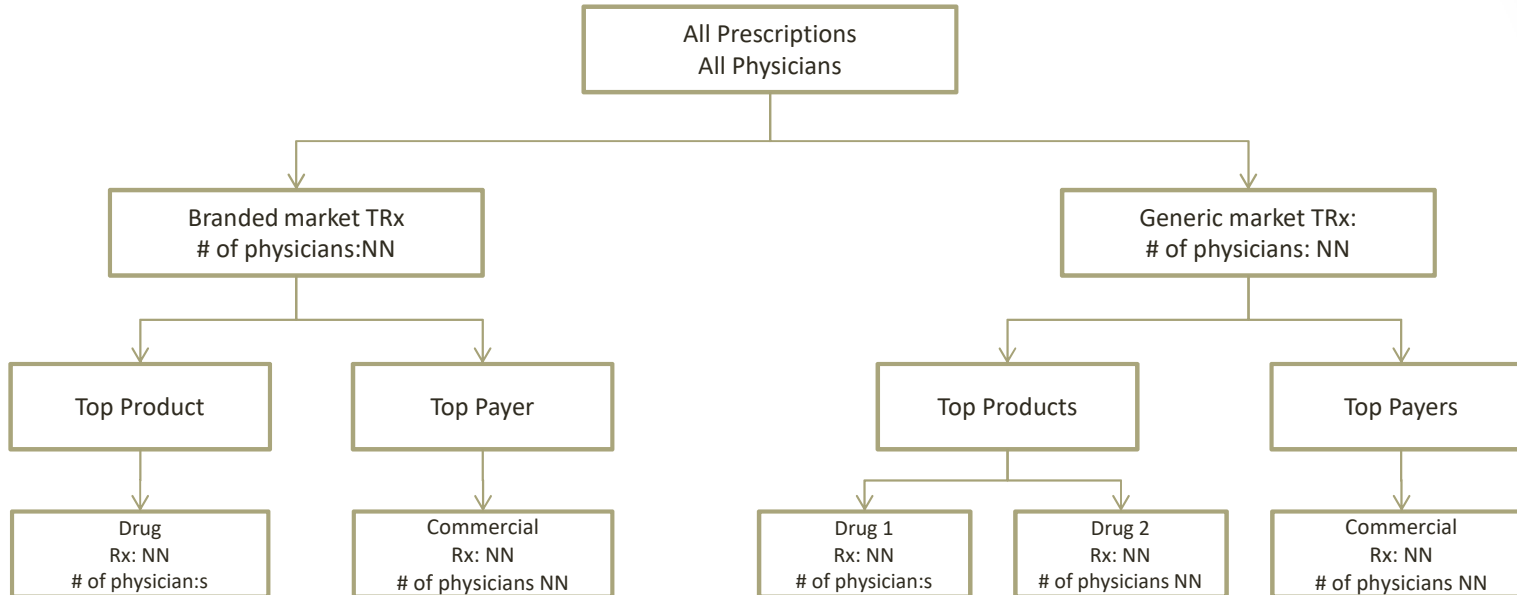
- Market: AED V Market
- Sales force: 2016Q3 Alignment
- Exponent PlanTrak: TRx, Call and Samples data
- Time period; YTD 2016

| Branded vs. Generic share comparison

| TRx share by payer group

- Cash, Medicaid, Third party, Commercial and Medicaid

Market Overview



Results

- Generated a segmented file of target prescribers to Pharma Clients derived from machine learning based on observed Drug prescribing behaviors
 - Refocused to include only those prescribers in “on-label” therapeutic areas
 - Overlaid with current identification of Pharma clients' active calling targets (those prescribers on Pharma clients' active target list – Target Flag=Y)
 - Includes NPI identifiers
- Conduct campaigns to test and measure different sales, promotional, marketing, communication and channel stimuli, in an abbreviated period of time in order to determine which stimuli have the most effect on migrating target prescriber script-writing behavior to higher levels of performance

Business Case Components – Summary

- Conclusion:
 - Pharma Clients can generate incremental Drug sales of \$452,000* annually in test market by refining the current prescribing segmentation to include machine learning recommendations to apply different promotional stimuli to different Prescribers segments
 - Applied nationally this would represent an annual incremental sales gain of \$15,912,000.

*Annual value of a NRX estimated at \$300. Lifetime Value not factored in

Business Impact Assumptions/Methodology

- Assumptions:
 - Target: Top 10% prescribers in therapeutic area, (branded affinity)
 - 50% of prescribers contacted will write 4-12 incremental Drug prescriptions per year
 - Value of a NRx prescription to Pharma Clients is \$300.00 annually illustrative (~\$500)
- Business Case
 - Test market: \$452,400 incremental revenues, year 1 from prescribers who modify their behavior
 - National: Top 10% of national branded affinity prescribers -- \$15,912,000

Business Case Assumptions

US Prescribers	1326000	1,326,000
15% in Target Therapeutic Area	0.1	132,600
Top 10%	0.1	13,260
50% will grow their prescribing volumes	0.5	6,630
by 4-12 new prescriptions per year (choose 8 as midpoin	8	53,040
Each new script is worth \$300 (nominal \$500)	\$300	\$15,912,000
Ten year lifetime value (no decay)	10	\$159,120,000

Conclusions

- A traditional statistical approach from the data provided and modeled to identify and segment high value / high potential prescribers, did not provide hoped-for results
- Machine Learning specifically identified Prescribers that have high potential and high propensity to write materially more Drug prescriptions than they do presently
- Only a small fraction of these currently “targeted” by Pharma Client, and most could be immediately tested by beginning promotional activities that represent high, immediate value and study validation for Pharma Clients
- In addition, market micro-segments will be investigated as other potential segments (Ex: transient ER Prescribers, out of territory...)

Second Application: Machine Learning Applied to New Product Launch Targeting

- The First 6 Months of a New Product Launch Determines the Lifetime Slope for the Duration of a Brand's Life Cycle
- Critical Factors to Launch Success
 - Who are My Prescribers and How to Stimulate Adoption Sooner Among Current Non-Prescribers that "look like" Early Adopters?
 - Is There Variance Between Anticipated Writers and Reality?
 - What Affinities Differentiate Early Adapters?
 - Which Brand Messages and Promotional Tactics/Stimuli are Most Effective?
 - Are My Resources Aligned for Maximum Effectiveness?
 - Am I able to Test and Respond in Near Real-Time?
 - Who are My Most Effective Sales Representatives, DSMs, RSDs
 - What Behaviors Differentiate Them From the Rest of the Pack?
- Target the Highest Propensity Prescribing Prospects based on Machine Learning approach, with the Highest Performing Promotional Stimuli
 - Measure, track and iterate

Accelerate Trial and Early Adoption by Physicians

- Quickly characterize early adopters and identify similar prescriber targets from those that are not writing yet
 - Track weekly and monthly longitudinally in sample test market
 - Accurately capture and track stimuli to determine which is working best and improve
 - Enhance prescribers with external demographics to accelerate targeting accuracy, updated at a minimum, monthly, preferably every 2 weeks
- Build, within 30 days, an evolved national target segmentation, with prioritized targets to accelerate Drug XX adoption and build share toward launch goals
- Deploy pragmatic test and learn to capitalize on the remaining 2016 launch window
- Track results retargeting segments that work

Rapid Analysis and Modeling of New Drug Sales While Targeting Promotional Stimuli To Grow Early Adoption

- How:
 - Capture prescriber sales of Drug weekly from IMS, apply additional external data enhancement characteristics to early adopting physicians and identify other physicians (that have been non-prescribing so far) for targeted promotional stimuli
 - Rapidly find non-prescribers that look statistically or mathematically like Drug XX early adopters, retargeting non-prescribers for alternative stimuli THE VERY NEXT WEEK
 - Machine Learning produces rapid results no need to wait for lengthy statistical analysis

New Drug Business Target Marketing Business Case

- Goal: Quadruple Pharma Clients New Drug Prescribers
- Assumed Baseline: 100 Drug XX Prescribers now, writing 4 NRx/month;\$300 per NRx
- Find 400 (+/-) similar in National Prescribing Population
 - 1,326,000 top .1% -- 1326 (Total Target Population expected to be much larger nationally)
 - Convert 1/3 of these – 442 to writing Drug XX scripts within the next 90 days**
 - Avg 4 NRX / month
 - \$300 annually net incremental to Pharma Clients per NRX
 - \$530,400 in 1st year incremental revenues (442*4*300)*
 - Market share gain of: 442 % (442/100)

*Lifetime value and decay curves not calculated

Data Mining Technologies Nuggets Machine Learning System

- *designed for today's large and/or complex datasets*
- *powerful pattern discovery engine which maps patterns into If – Then rules*
- *employs proprietary machine learning search through space of all possible rules*
- *can model thousands of predictor variables*
- *automatically builds, validates, and tests models against hold-out data*
- *rapid model development - short learning curve*
- *auditable results understandable by management*
- *builds predictive and descriptive models*
- *pattern drill -down capabilities for exploratory analysis*
- *rules properties such as confidence factor, coverage, and statistical significance*
- *handles missing values of predictor variable without imputation*
- *non statistical methodology which permits correlations among predictor variables*
- *does not require predictor variables to be independent of each other*

Modeling with Nuggets[®]

- *Unique Automatic one step model build and test: validates results and applies predictions to new data*
- *Has been used to build predictive and descriptive models, optimization models, classifier models, ROI models, scoring models*
- *No Neural nets, Tree Builders or statistics needed*
- *Minimizes potential for over fitting*
- *K- fold or jackknife validation*
- *Several built-in binning methods (equal height, equal width, normal, custom)*
- *Attribute significance analysis*
- *Gains chart, Lift chart, Confusion matrix*

Data Analysis with Nuggets[®]

- *Examines data for significant relationships, including Categorical data analysis, and Numerical data analysis and mixed analysis*
- *Attribute analysis for correlations requires no parametric assumptions*
- *Normative and predictive analysis*
- *Attribute importance*
- *Attribute reduction (significance selectable)*
- *Rule Significance test (significance selectable)*
- *Supervised cluster analysis, Supervised segmentation*
- *Implicit automatic imputation for missing values*

Data Visualization and Presentation with Nuggets[®]

- *Graphs and charts of attribute distributions*
- *Graphical progress of model build progress*
- *Graphical views of data values distribution*
- *Visual comparison of competing model results*
- *Report export for customized presentations*
- *Create spreadsheets for customized analysis of results*

For Questions or More Information about
this Presentation or about Nuggets[®]

Please Contact :

Dr. Michael Gilman

at Data Mining Technologies Inc.

mgilman@data-mine.com