ADVANCED PROGRAM IN AI – POWERED MARKETING

AI-Based MROI Optimization for Tally

Capstone Project

Tally Solutions Pvt. Ltd.

● .Incorporated in late 1991 - India's leading business management software product company.

● Imports & exports computer software programs for use in Industries, Banks, Financial Institutions, Hotels and other Commercial establishments.

● Offers a wide range of products: Tally Prime, Tally Primer Server & Shopper 9.

● Offers services like ERP implementation, Training on product usage, Deployment, Software assurance, Data migration, Business advisory, and Integration.

● Offers custom built, business specific, country specific solutions, data migration tools, device integrations, 3rd party integration, and add-ons.

● Ticket size: Rs. 600/month - Rs. 54,000/perpetual

Google Analytics Data

12/10/20-19/10/21

Data Variables

Audience

1. Client ID
2. User Type
3. Gender
4. Age
5. In-Market Segment
6. Browser
7. Device Category

Behavior

1. Bounce Rate
2. Page/Session
3. Avg. Session Duration
4. Time on Page
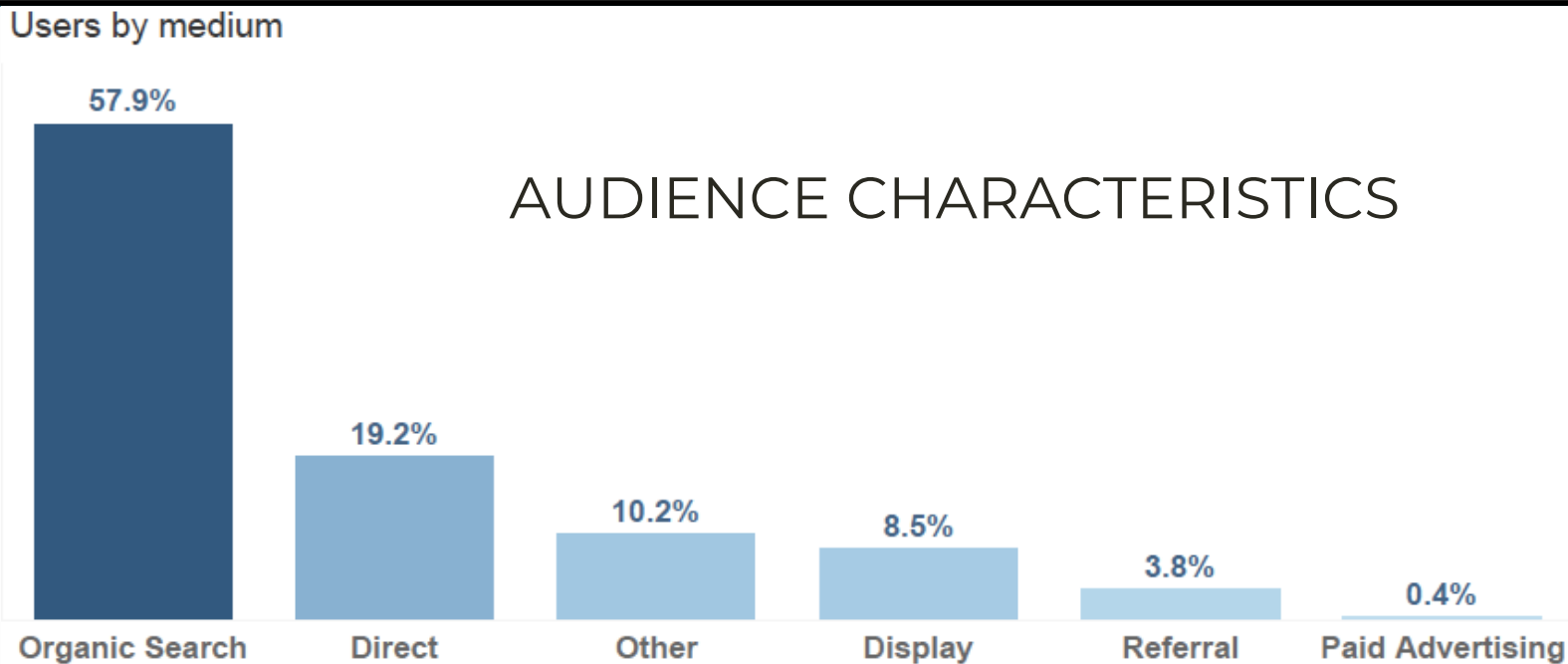5. Unique Page Views
6. Page Load Time ms
7. Page Download Time ms

Acquisition

1. Source
2. Medium
3. Channel
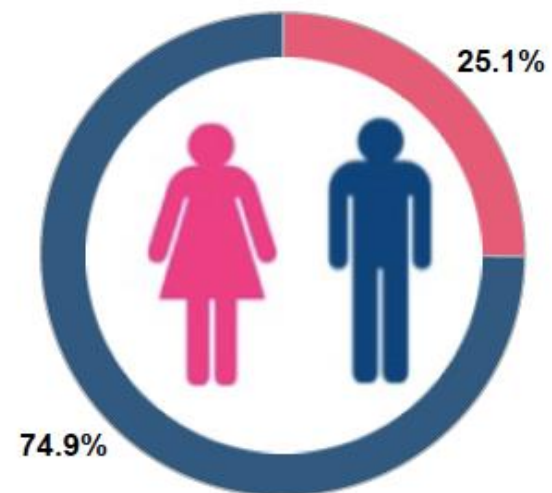4. Sessions
5. % New Sessions
6. Referrals

Conversion

1. Avg. Document Content Loaded Time sec
2. Goal Completions
3. Goal Conversion Rate
4. Avg. Order Value
5. Transactions
6. Revenue
7. Ecommerce Conversion Rate

AUDIENCE CHARACTERISTICS

Users by medium

- Organic Search: 57.9%
- Direct: 19.2%
- Other: 10.2%
- Display: 8.5%
- Referral: 3.8%
- Paid Advertising: 0.4%

Users by gender

- 25.1%
- 74.9%

Users by age

- 18-24: 25.4%
- 25-34: 30.4%
- 35-44: 19.4%
- 45-54: 12.0%
- 55-64: 7.8%
- 65+: 5.0%

Users by Country

- United States
- India
- United Arab Emirates
- Other

- 0.5%
- 3.5%
- 0.7%
- 95.3%
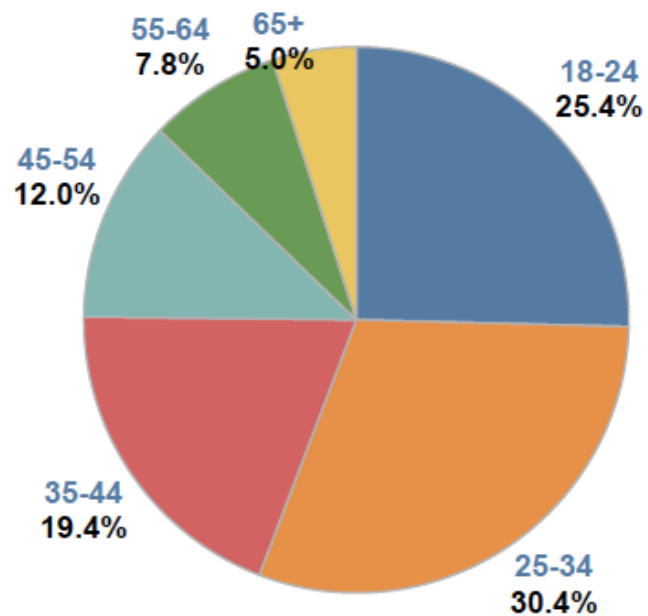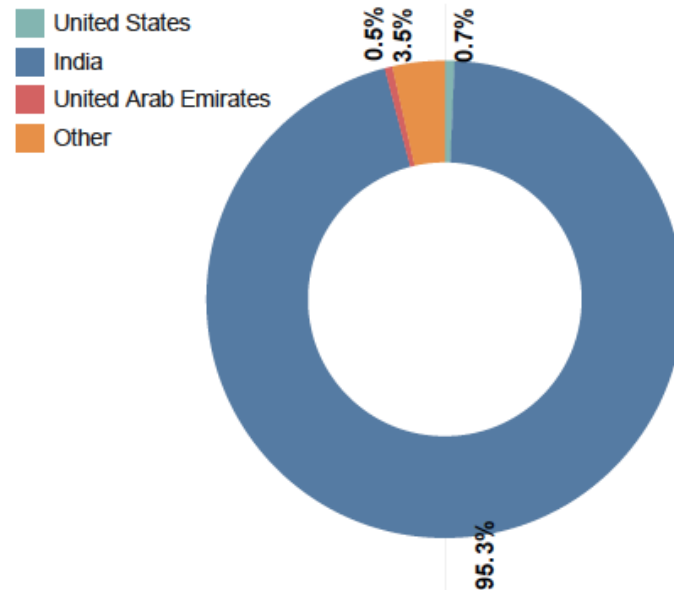
Users by In-Market Segment

- Others: 50,100
- 40,010
- Education/Test Preparation &: 40,509
- 21,164
- Financial: 51,077
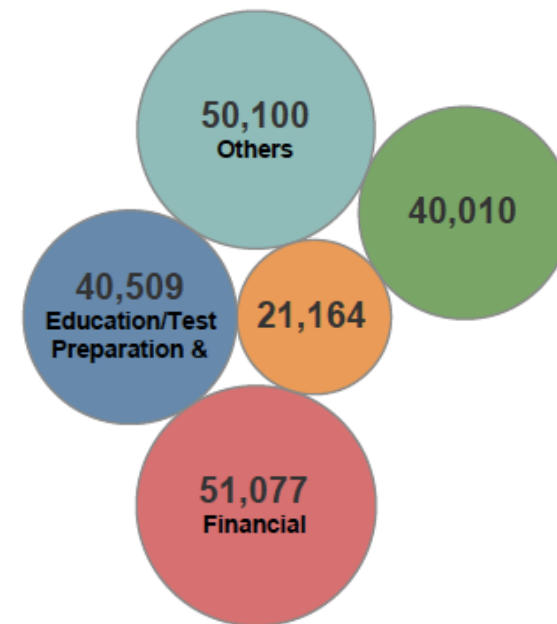
Project Objectives

To effectively predict the likelihood of conversion of website users.

Objective 1

To effectively predict the potential amount of revenue generation by a converted user.

Objective 2

To predict the propensity score for a lead for more effective marketing and sales conversions.

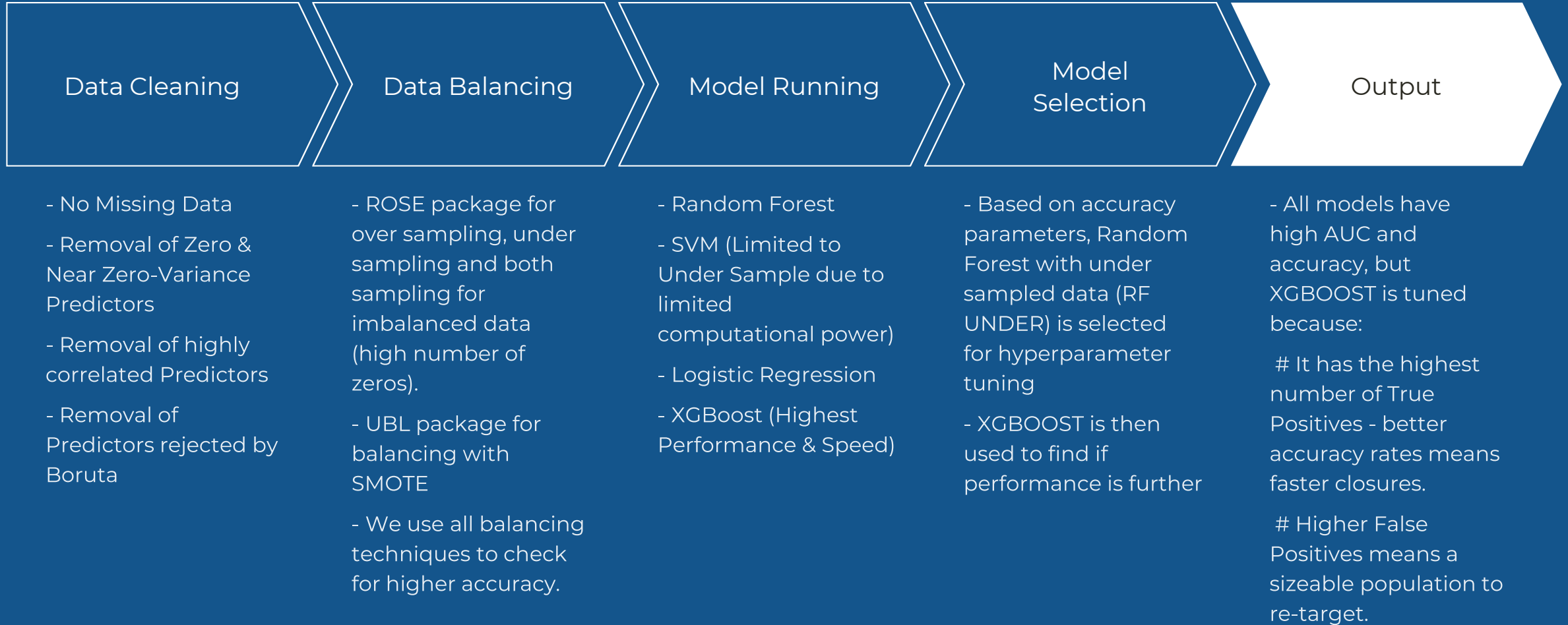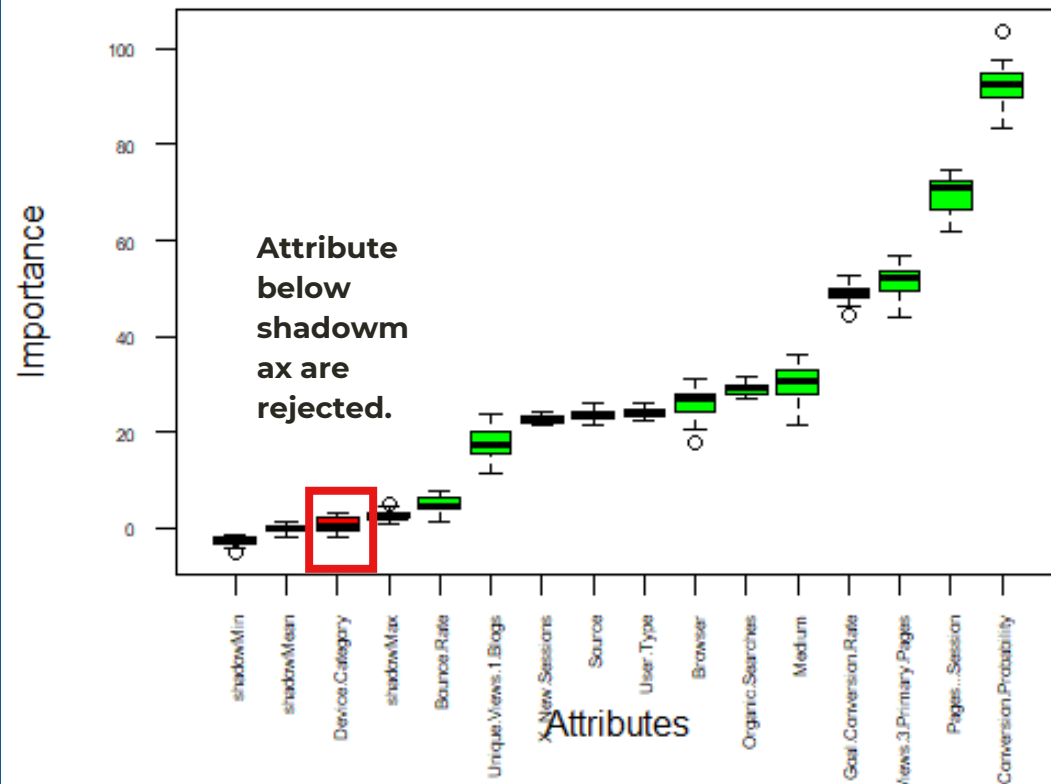Objective 3

Objective 1

Website User Conversion Prediction

Model Building Process

Data Cleaning	Data Balancing	Model Running	Model Selection	Output

Data Cleaning
- No Missing Data
- Removal of Zero & Near Zero-Variance Predictors
- Removal of highly correlated Predictors
- Removal of Predictors rejected by Boruta

Data Balancing
- ROSE package for over sampling, under sampling and both sampling for imbalanced data (high number of zeros).
- UBL package for balancing with SMOTE
- We use all balancing techniques to check for higher accuracy.

Model Running
- Random Forest
- SVM (Limited to Under Sample due to limited computational power)
- Logistic Regression
- XGBoost (Highest Performance & Speed)

Model Selection
- Based on accuracy parameters, Random Forest with under sampled data (RF UNDER) is selected for hyperparameter tuning
- XGBOOST is then used to find if performance is further

Output
- All models have high AUC and accuracy, but XGBOOST is tuned because:
  # It has the highest number of True Positives - better accuracy rates means faster closures.
  # Higher False Positives means a sizeable population to re-target.

Boruta Output

Attribute below shadowmax are rejected.

```
> print(boruta)
Boruta performed 25 iterations in 40.82815 mins.
 12 attributes confirmed important: Avg....Conversion.Probability, Bounce.Rate, Browser,
Goal.Conversion.Rate, Medium and 7 more;
 1 attributes confirmed unimportant: Device.Category;
```

```
> #Tentative Fix
> #bor = TentativeRoughFix(boruta) - no Tentative attributes
> #print(bor)
> attStats(boruta)
```

	meanImp	medianImp	minImp	maxImp	normHits	decision
Source	23.6568111	23.4915407	21.422067	26.070181	1.00	Confirmed
Medium	30.0750795	30.6790965	21.467105	36.099959	1.00	Confirmed
Browser	26.1972574	27.0330700	17.988811	31.004423	1.00	Confirmed
Device.Category	0.5775765	0.3650608	-1.785668	3.085533	0.08	Rejected
User.Type	24.2863212	24.3265587	22.476776	25.893044	1.00	Confirmed
Goal.Conversion.Rate	48.7786149	48.8075049	44.402067	52.632598	1.00	Confirmed
Avg....Conversion.Probability	92.0509342	92.6483118	83.563333	103.776009	1.00	Confirmed
X..New.Sessions	22.6344388	22.6391100	21.431565	24.275372	1.00	Confirmed
Organic.Searches	28.9305978	29.2526275	26.936989	31.377794	1.00	Confirmed
Bounce.Rate	4.6836014	4.6343437	1.024198	7.809514	0.84	Confirmed
Pages...Session	70.0586755	71.1618694	61.865961	74.683473	1.00	Confirmed
Unique.Views.1.Blogs	17.6272767	17.2055389	11.349029	23.578610	1.00	Confirmed
Unique.Views.3.Primary.Pages	51.4593525	52.0602702	43.973884	56.975048	1.00	Confirmed

Prediction Models

	RF UNDER		RF OVER		RF BOTH		RF SMOTE			SVM UNDER		LOGISTIC OVER		LOGISTIC UNDER		LOGISTIC BOTH		LOGISTIC SMOTE	
	0	1	0	1	0	1	0	1		0	1	0	1	0	1	0	1	0	1
0	34049	50	34947	80	35019	70	36765	100		30489	39	32600	61	32645	66	32570	62	32717	63
1	6049	424	5151	394	5079	404	3333	374		9609	435	7498	413	7453	408	7528	412	7381	411
Accuracy	0.8497		0.8711		0.8731		0.9154			0.7622		0.8137		0.8147		0.8129		0.8165	
Sensitivity	0.8491		0.87154		0.87334		0.9169			0.76036		0.81301		0.8141		0.81226		0.81593	
Specificity	0.8945		0.83122		0.85232		0.789			0.91772		0.87131		0.8608		0.8692		0.86709	
AUC	0.872		0.851		0.863		0.853			0.839		0.842		0.837		0.878		0.842	

- Fairly high AUC
- Higher True Positives: Better accuracy rate, faster closure
- Higher False Positives: A sizeable population to target

	RF UNDER TUNED	
	0	1
0	34634	56
1	5464	418
Accuracy	0.8639	
Sensitivity	0.86373	
Specificity	0.88186	
AUC	0.873	

	XGBOOST UNDER		XGBOOST UNDER (TUNED)		XGBOOST SMOTE		XGBOOST SMOTE (TUNED)	
	0	1	0	1	0	1	0	1
0	33313	45	32260	34	34465	51	33880	44
1	6785	429	7838	440	5633	423	6218	430
Accuracy	0.8317		0.806		0.8599		0.8457	
Sensitivity	0.83079		0.80453		0.85952		0.84493	
Specificity	0.90506		0.92827		0.89241		0.90717	
AUC	0.868		0.866		0.876		0.876	

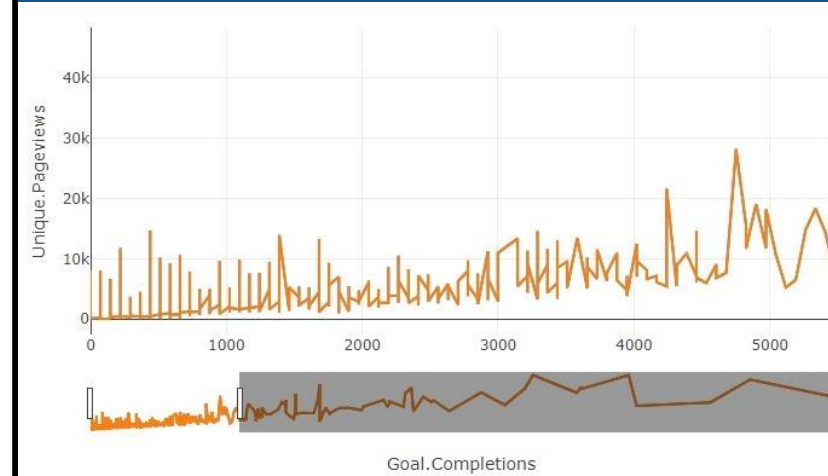# Significant Attributes

1 Pages/Session

2 Unique Views - 3 Primary (Product) Page
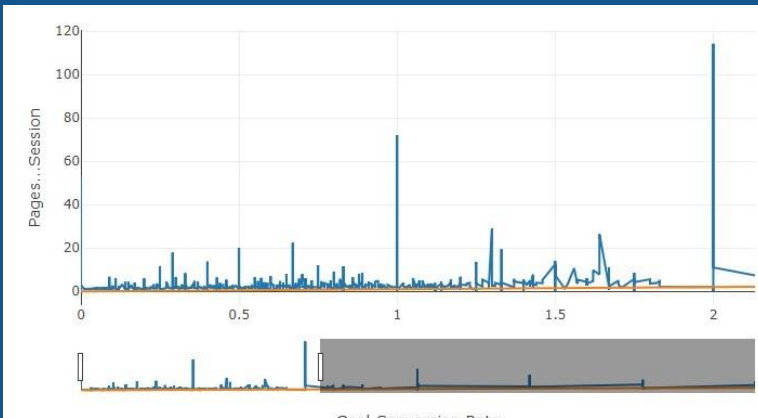
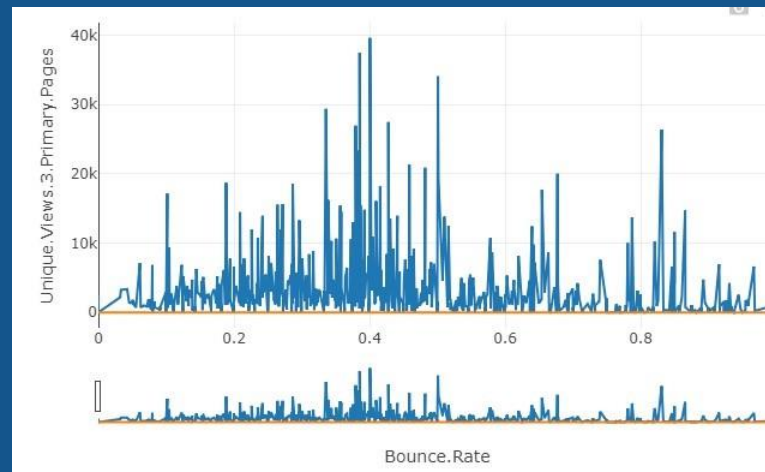3 Goal Conversion Rate

4 Bounce Rate

Conversion Visualization

- Page views and sessions taper after a point

-This means that those who do not convert in the first instance, lose their way through the journey

- NEED – more engaging content to minimize drop off. Better UX
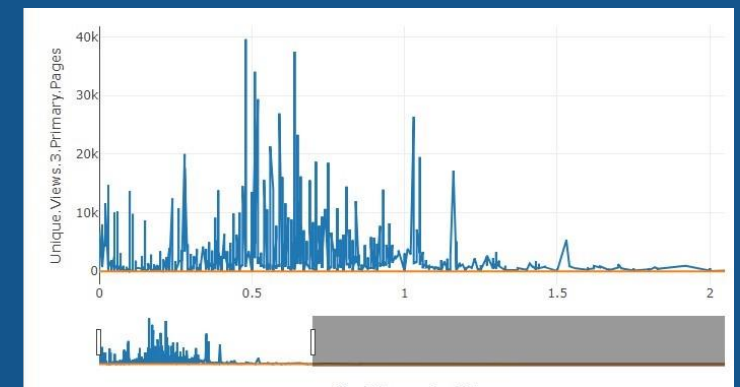

PRODUCT PAGES UNIQUE VIEW VS GOAL CONVERSION

PAGE SESSION VS GOAL CONVERSION

PRODUCT PAGES UNIQUE VIEW VS BOUNCE RATE

UNIQUE PAGE VIEW VS GOAL COMPLETION

Potential Customer Revenue Generation Prediction

Model Building & Selection Process

| Data Cleaning | Model Running | Model Selection | Output |

Data Cleaning

- No Missing Data
- Removal of Zero & Near Zero-Variance Predictors
- Removal of highly correlated Predictors
- Removal of Predictors rejected by Boruta

Model Running

- Random Forest
- SVM
- Linear Regression

Model Selection

- Based on lowest RMSE, Random Forest is selected.
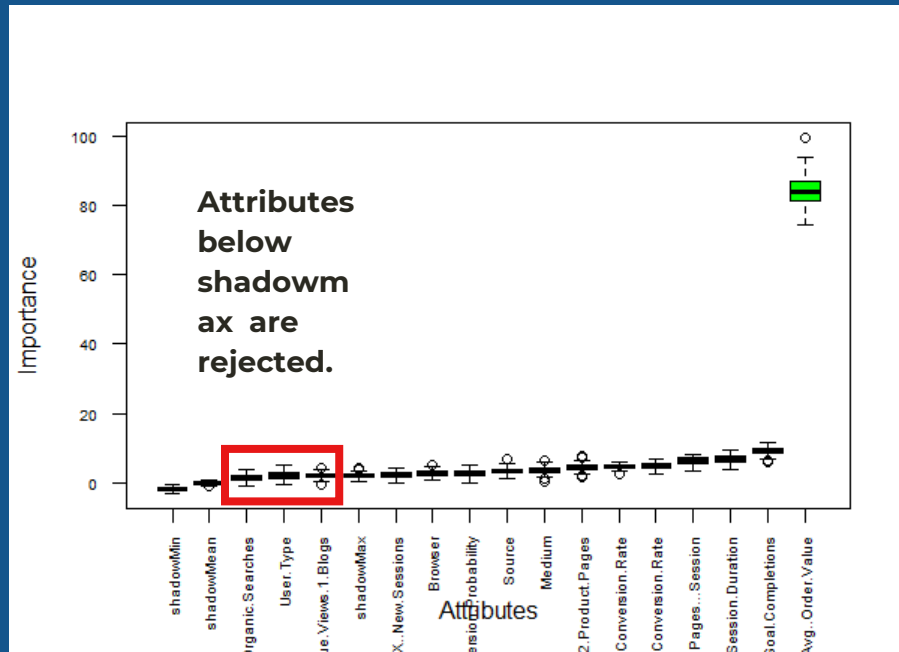- Random Forest model is tuned for hyperparameters

Output

- Tuned Random Forest model has the lowest RMSE, thus is the selected model.
- mtry=13; ntree=1078; nodesize=17

	Linear Regression	SVM	SVM Tuned	Random Forest	Random Forest Tuned
RMSE	188,906.8	169,991.8	188,268.8	131,298.8	128,416.4

Lowest RMSE (in Rs.)

Boruta Output

Attributes below shadowmax are rejected.

```
> print(boruta)
Boruta performed 99 iterations in 1.029329 mins.
 11 attributes confirmed important: Avg....Conversion.Probability, Avg..Order.Value,
Avg..Session.Duration, Browser, Ecommerce.Conversion.Rate and 6 more;
 1 attributes confirmed unimportant: Organic.Searches;
 3 tentative attributes left: Unique.Views.1.Blogs, User.Type, X..New.Sessions;
```

```
> bor = TentativeRoughFix(boruta)
> print(bor)
Boruta performed 99 iterations in 1.029329 mins.
Tentatives roughfixed over the last 99 iterations.
 14 attributes confirmed important: Avg....Conversion.Probability,
Avg..Order.Value, Avg..Session.Duration, Browser,
Ecommerce.Conversion.Rate and 9 more;
 1 attributes confirmed unimportant: Organic.Searches;
```

```
> attStats(boruta)
```

	meanImp	medianImp	minImp	maxImp	normHits	decision
Source	3.315803	3.312399	1.22386745	6.600852	0.8787879	Confirmed
Medium	3.627556	3.585982	0.33361754	6.527163	0.8686869	Confirmed
Browser	2.768842	2.715080	0.82013418	5.058436	0.7777778	Confirmed
User.Type	1.985728	1.874640	-0.35309445	5.045240	0.5050505	Tentative
Goal.Completions	9.196945	9.382560	5.99980111	11.449246	1.0000000	Confirmed
Goal.Conversion.Rate	4.878319	4.961767	2.63005718	6.679446	1.0000000	Confirmed
Avg....Conversion.Probability	2.657553	2.815063	0.05277964	5.156683	0.6666667	Confirmed
X..New.Sessions	2.186498	2.236503	-0.13690906	4.157938	0.5151515	Tentative
Organic.Searches	1.432711	1.640319	-1.13135613	3.609136	0.1616162	Rejected
Pages...Session	6.336342	6.564723	3.37234135	8.273775	1.0000000	Confirmed
Avg..Session.Duration	6.738988	6.788661	3.93238629	9.404504	1.0000000	Confirmed
Ecommerce.Conversion.Rate	4.603161	4.634112	2.36259417	5.897567	0.9696970	Confirmed
Unique.Views.1.Blogs	2.052731	2.053584	-0.34084278	4.379965	0.4949495	Tentative
Unique.Views.2.Product.Pages	4.600019	4.601171	1.65400541	7.529916	1.0000000	Confirmed
Avg..Order.Value	84.229719	83.861630	74.31424801	99.574055	1.0000000	Confirmed

Significant Attributes

1 Avg. Order Value

2 Avg. Session Duration

3 Goal Completions

4 Ecommerce Conversion Rate
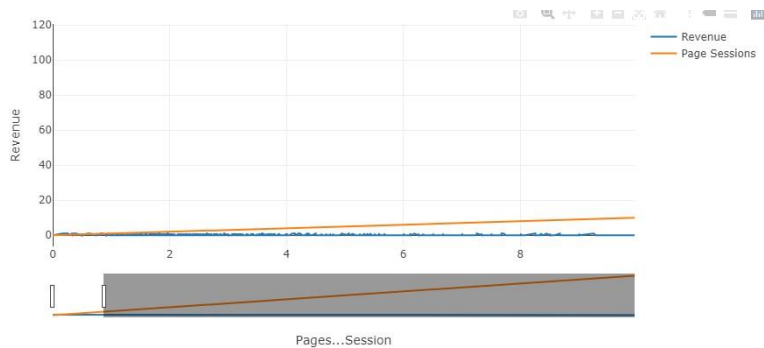
REVENUE VS PAGE SESSIONS
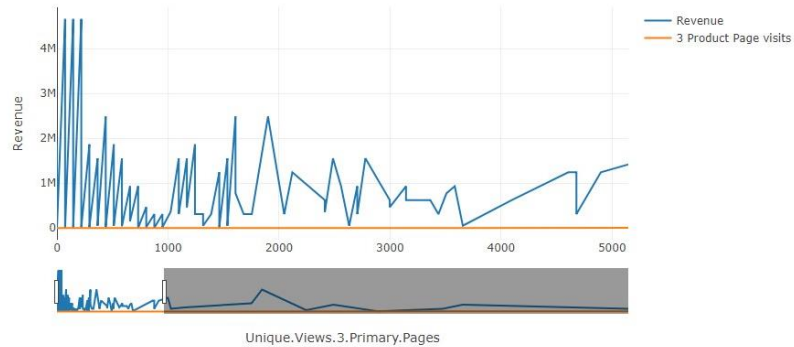
REVENUE VS UNIQUE PAGE VIEWS (PRODUCT PAGE)

REVENUE VS GOAL CONVERSION RATE

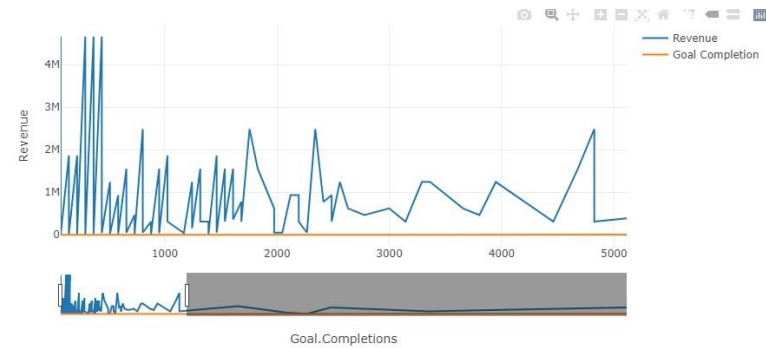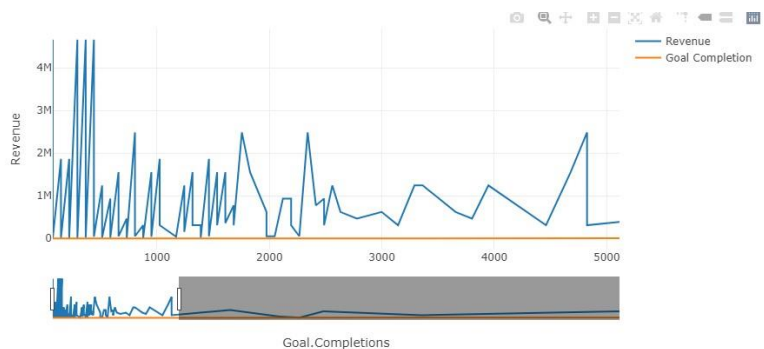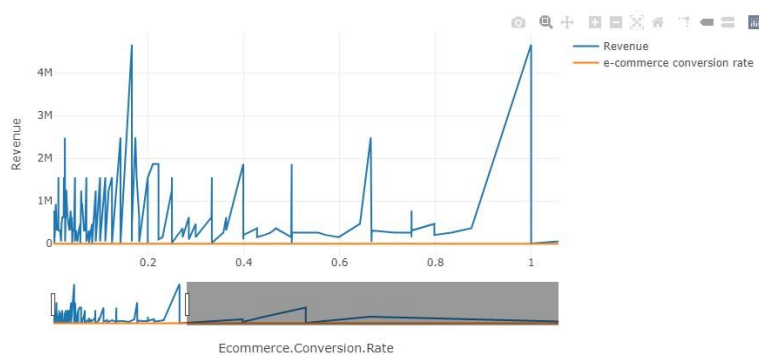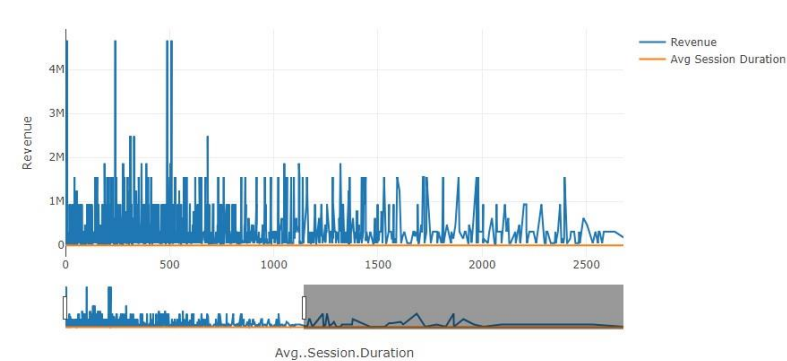Observation: Revenue is consistent up to a particular point, after which it becomes erratic and sparse.

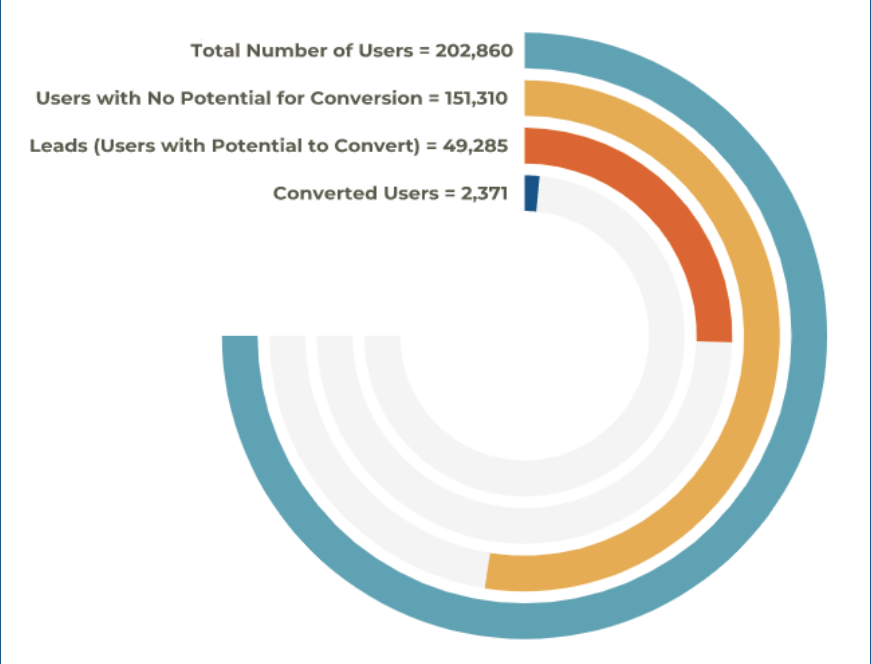REVENUE VS GOAL COMPLETION

REVENUE VS ECOMMERCE CONVERSION RATE

REVENUE VS AVERAGE SESSION DURATION

Objective 3

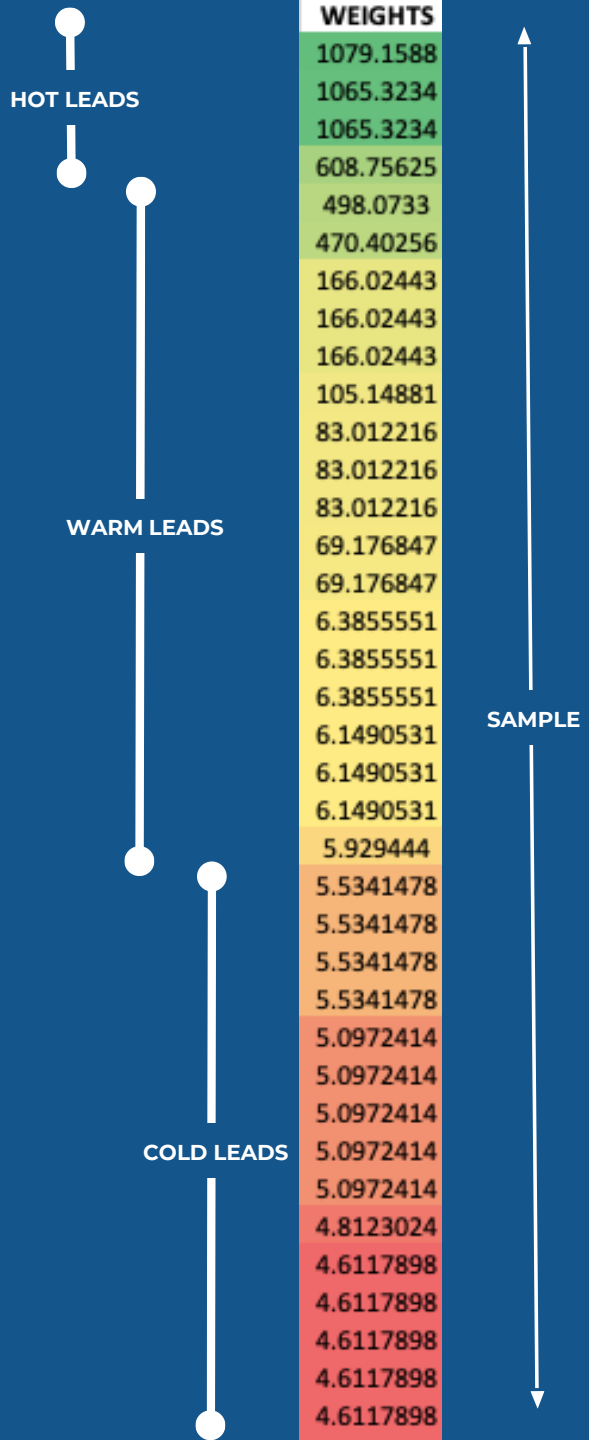# Lead Propensity Scoring for Improved MROI

Lead Propensity Scoring using Greedy Matching

- User Data Breakdown

Total Number of Users = 202,860

Users with No Potential for Conversion = 151,310

Leads (Users with Potential to Convert) = 49,285

Converted Users = 2,371

Data Snapshot

Client ID	Source	Medium	Browser	User.Type	Goal.Convers	X..New.Sessi	Organic.Sear	Bounce.Rate	Pages...Sessi	Unique.View	Unique.View	Revenue	weights	subclass
28164	direct	none	Edge	New Visitor	0.8772345	0.6403005	-0.871571	-1.030427	-0.94718	-0.441915	-0.158325	0	1079.1588	402
13706	direct	none	Chrome	New Visitor	0.8772345	0.6403005	-0.871571	-1.030427	-0.94718	-0.441915	-0.158325	0	1065.3234	314
67637	direct	none	Chrome	New Visitor	0.8772345	0.6403005	-0.871571	-1.030427	-0.94718	-0.441915	-0.158325	0	1065.3234	314
159045	direct	none	Edge	Returning Vis	0.8772345	-1.560417	-0.871571	-1.030427	-0.94718	-0.441915	-0.158325	0	608.75625	602
140036	direct	none	Internet Expl	Returning Vis	2.4632714	-1.560417	-0.871571	-1.030427	-0.94718	-0.441915	-0.158325	0	498.0733	580
47	direct	none	Internet Expl	Returning Vis	0.8772345	-1.560417	-0.871571	-1.030427	-0.94718	-0.441915	-0.158325	0	470.40256	7
34157	google	organic	Chrome	Returning Vis	0.8772345	-1.560417	-0.871571	-1.030427	-0.94718	-0.441915	-0.158325	0	166.02443	424
72070	Others	referral	Chrome	Returning Vis	0.8772345	-1.560417	-0.871571	-1.030427	-0.94718	-0.441915	-0.158325	0	166.02443	517
113210	google	organic	Chrome	Returning Vis	0.8772345	-1.560417	-0.871571	-1.030427	-0.94718	-0.441915	-0.158325	0	166.02443	424
154023	google	organic	Chrome	Returning Vis	0.8772345	-1.560417	-0.871571	-1.030427	-0.94718	-0.441915	-0.158325	0	166.02443	424
181751	google	organic	Chrome	Returning Vis	0.8772345	-1.560417	-0.871571	-1.030427	-0.94718	-0.441915	-0.158325	0	166.02443	424
4808	direct	none	Internet Expl	New Visitor	0.8772345	0.6403005	-0.871571	-1.030427	-0.42425	-0.441915	-0.158325	0	138.35369	196
69730	direct	none	Internet Expl	New Visitor	0.8772345	0.6403005	-0.871571	-1.030427	-0.42425	-0.441915	-0.158325	0	138.35369	196
19113	direct	none	Chrome	Returning Vis	0.8772345	-1.560417	-0.871571	-1.030427	-0.94718	-0.441915	-0.158325	0	105.14881	351
22461	direct	none	Chrome	Returning Vis	0.8772345	-1.560417	-0.871571	-1.030427	-0.94718	-0.441915	-0.158325	0	105.14881	351
134218	direct	none	Chrome	Returning Vis	0.8772345	-1.560417	-0.871571	-1.030427	-0.94718	-0.441915	-0.158325	0	105.14881	351
141884	direct	none	Chrome	Returning Vis	0.8772345	-1.560417	-0.871571	-1.030427	-0.94718	-0.441915	-0.158325	0	105.14881	351
166101	direct	none	Chrome	Returning Vis	0.8772345	-1.560417	-0.871571	-1.030427	-0.94718	-0.441915	-0.158325	0	105.14881	351
36775	google	organic	Chrome	Returning Vis	2.4632714	-1.560417	-0.871571	-1.030427	-0.42425	-0.441915	-0.158325	0	83.012216	434
89341	direct	none	Internet Expl	Returning Vis	0.8772345	-1.560417	-0.871571	-1.030427	-0.42425	-0.441915	0.2662596	0	83.012216	540

HOT LEADS

WARM LEADS

COLD LEADS

SAMPLE

WEIGHTS
1079.1588
1065.3234
1065.3234
608.75625
498.0733
470.40256
166.02443
166.02443
166.02443
105.14881
83.012216
83.012216
83.012216
69.176847
69.176847
6.3855551
6.3855551
6.3855551
6.1490531
6.1490531
6.1490531
5.929444
5.5341478
5.5341478
5.5341478
5.5341478
5.0972414
5.0972414
5.0972414
5.0972414
5.0972414
4.8123024
4.6117898
4.6117898
4.6117898
4.6117898
4.6117898

Recommendations

Tally Web Analysis

	Keywords Identification and Inclusion	SEO Friendly Navigation	Correct and Compelling Content	Improve UX
https://tallysolutions.com/	7	9	8	6
https://tallysolutions.com/tally-prime/	7	9	9	6
https://tallysolutions.com/tally-prime-server/	7	9	9	6
https://tallysolutions.com/tally-shoper-9/	7	9	9	6
https://tallysolutions.com/download/	2	2	2	5
https://tallysolutions.com/buy-tally/	NA	9	6	8

MROI Optimization

IMPACT OF RECOMMENDATIONS

Attributes ————————— Propensity Outcomes

Recommendation	Unique Views - 3 Primary (Product) Page	Goal Conversion Rate	Bounce Rate	Avg. Order Value	Goal Completions	Pages Per Session	Avg. Session Duration	Ecommerce Conversion Rate	Cold Leads	Warm Leads	Hot Leads
Drip email campaigns for all 3 products consisting of interlinked, personalised, gated content that provides fresh information for continuous engagement and conversions in terms of form fills.		✓			✓				✓		
Enhanced, interactive and personalised user journey experience on the website through constantly fresh, dynamic, optimised and relevant content for all 3 products, including testimonials, use cases and videos.	✓		✓			✓	✓				
Improve SEO score through keywords optimization and quality backlinks from reputable publications via PR activities.			✓								
Value offers and discounts for increased conversion potential and higher order value, especially for subscription plans.		✓		✓	✓			✓		✓	
Direct Connect via Outbound Calling by Pre-sales team.											✓

Thank You