

# Customer Churn Prediction in Telecommunications

This presentation outlines a comprehensive analysis of customer churn in the telecommunications industry, leveraging machine learning to predict and prevent customer attrition.

 by **wanjiru Kinyanjui**

# Project Criteria (CRISP-DM)

## Business Understanding

Define the business problem and goals

## Data Understanding

Explore and understand the data

## Data Preparation

Clean, transform, and prepare the data

## Modeling

Build and train machine learning models

# Business Understanding

## Project Overview

Churn occurs when customers switch providers. Key factors: dissatisfaction, competitor pricing. Impact: lost revenue, reduced growth

## Business Problem

Identify customers at risk of churn. Develop strategies to retain them

# Data Description

# State

## Customer's geographic location

## Account Length

## Duration of the customer relationship

## International Plan

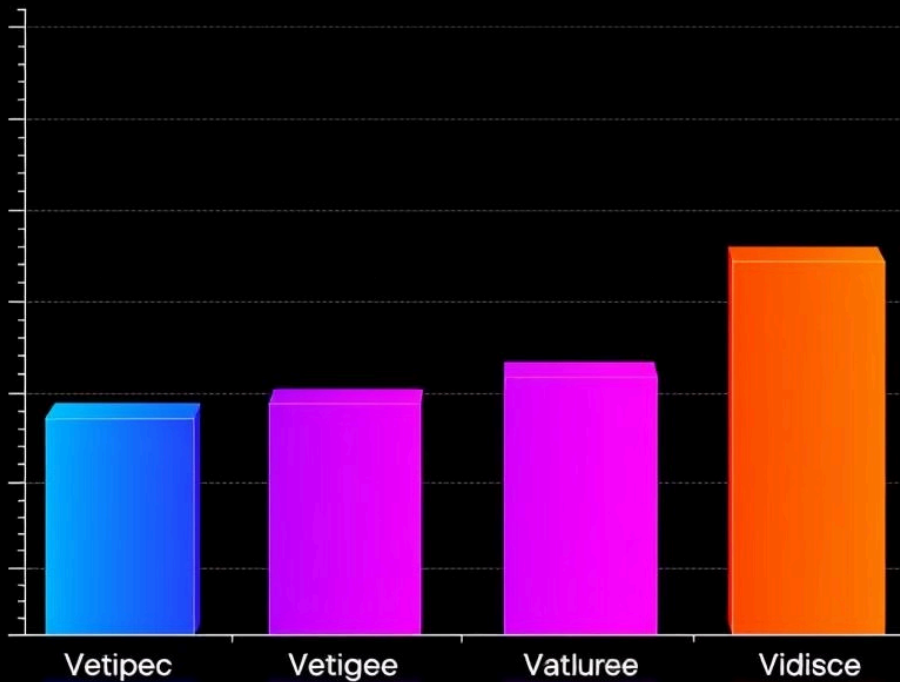
## Subscription to international calling

## Total Day Minutes

## Daytime call duration

Comprehensive Financial Performance Analysis - Q3 2024								
Product Line Overview			Market	Control	Current	Value	Revenue	Profit
Category			DATA1	DATA2	DATA3	DATA4	DATA5	DATA6
ID	Product	Segment	Revenue	Units	Price	Cost	Margin	Profit
1	20D4	Customers	\$6440	15120	52290	50566	\$6.00	4250
2	20D4	Customers	\$8140	11107	35999	54191	36.77	9380
3	20D4	Customers	\$6500	16170	32250	\$6390	\$6.00	5530
4	20D4	Customers	\$4100	16160	33150	57297	\$6.00	4350
5	20D4	Customers	\$6120	10180	56295	53170	50.00	4500
7	20D4	Customers	\$4130	15550	30299	34302	36.70	6571
8	20D4	Customers	\$6330	19990	38522	24390	35.00	3661
9	20D4	Customers	\$4140	17790	13225	56746	39.70	4500
6	20D4	Customers	\$4140	17359	32400	36197	32.42	5599
7	20D4	Customers	\$0.45	10990	19505	28193	16.42	2580
8	20D4	Customers	\$9106	10109	33396	36297	11.00	3360
8	20D4	Customers	\$0.06	19590	35396	00171	11.00	9380
9	20D4	Customers	\$0.60	10162	25599	96205	10.00	4560
5	20D4	Customers	\$6100	11770	06500	00100	\$6.00	0387
9	20D4	Customers	\$6440	10170	32995	06797	39.70	3070
6	20D4	Customers	\$6140	19309	32298	04292	\$2.77	6471
7	20D4	Customers	\$6348	14594	18399	04757	42.70	\$277
7	20D4	Customers	\$4.40	16740	10560	04270	42.70	5560
8	20D4	Customers	\$6160	16155	53088	06250	11.47	0560
9	20D4	Customers	\$6165	10750	38169	03257	\$6.75	0550
0	20D4	Customers	\$6150	16258	20100	54217	\$1.70	0500
6	20D4	Customers	\$6170	10790	08996	04440	\$2.00	0560
7	20D4	Customers	\$4170	16355	36560	06726	\$2.00	0560
8	20D2	Customers	46340	13050	23160	15356	11.10	0560
7	20D4	Customers	\$6310	16990	32550	08199	\$1.10	0387
6	20D4	Customers	\$6540	15370	22790	06795	\$4.00	0541
7	30D4	Customers	\$4190	19577	32596	16297	25.00	4542
3	20D4	Customers	\$6336	10350	39400	00100	\$6.00	0387
6	30D4	Customers	\$4120	10399	33194	24395	34.72	3440

# Subjective Model Performance



## Evaluation Metrics



### Accuracy

Overall model performance



### Precision

Correct positive predictions



### Recall

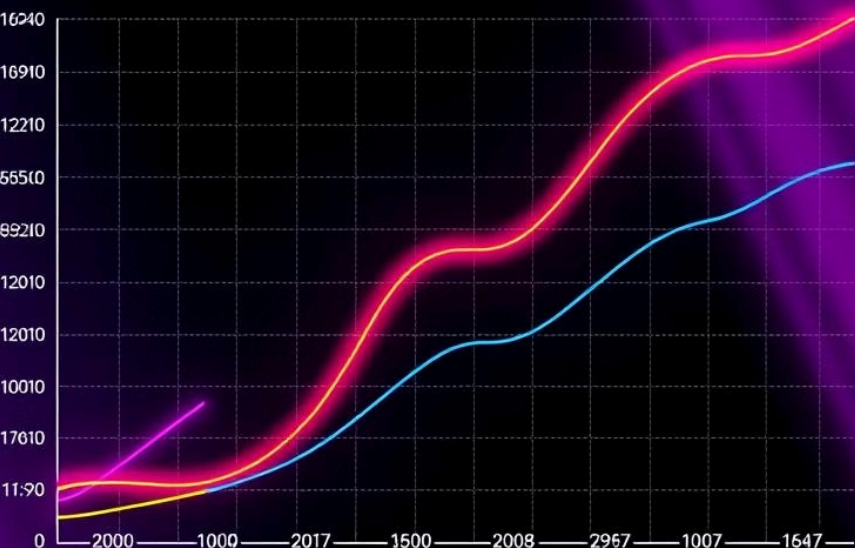
Correctly identified churn cases



### F1 Score

Balance of precision and recall

Model performance comparison  
metrics and trends



## Model Comparison - Accuracy

93.7%

Random Forest

Highest accuracy

91.75%

Decision Tree

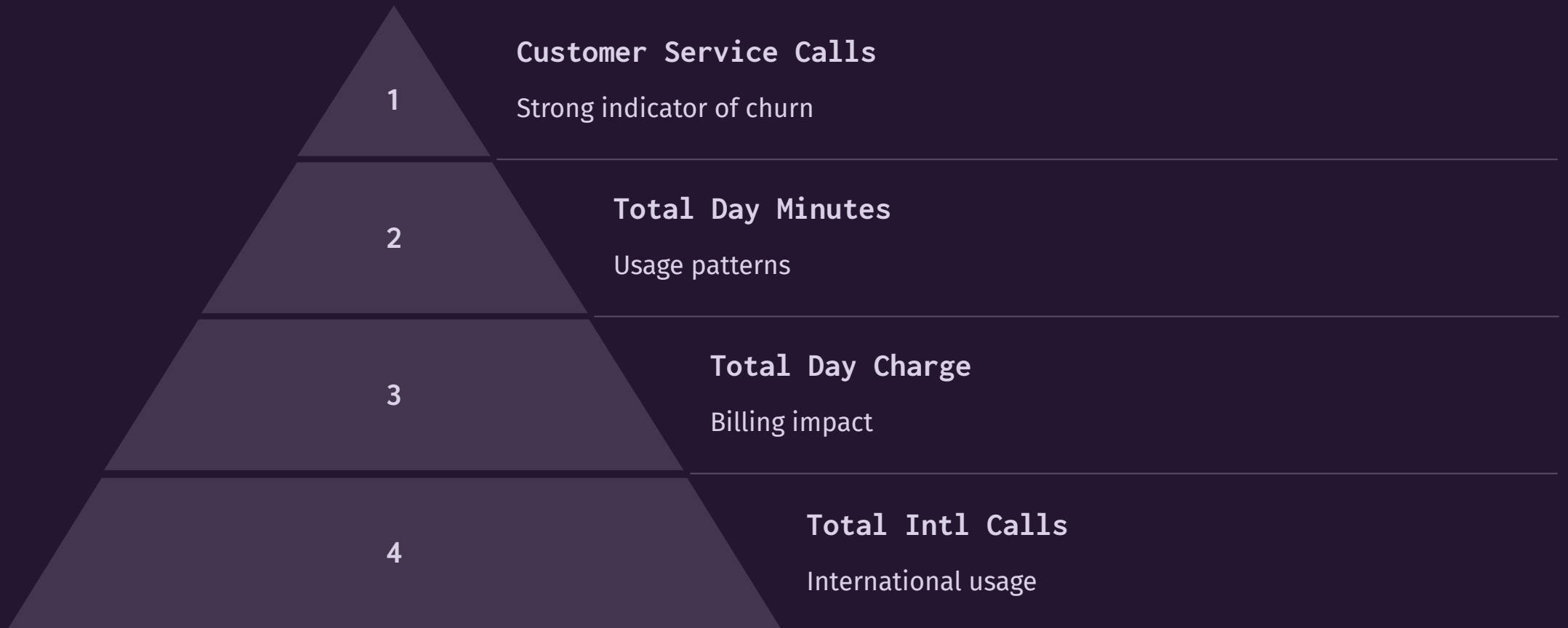
Strong performance

86.21%

Logistic Regression

Baseline model

# Key Insights from Feature Importance



# Recommendations for the Business

1

## Improve Customer Service

Reduce call volume, enhance service quality

---

2

## Evaluate Pricing Structure

Adjust charges based on usage

---

3

## Engage with High-Risk Customers

Reach out to customers with high usage