# **Magic Position AI - Proposal**

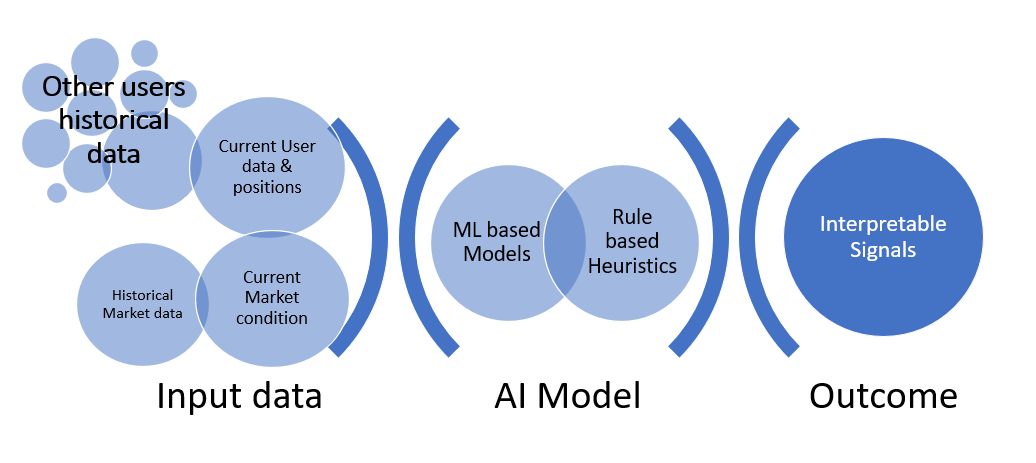
Today’s customers live in a world where information is fired at them day in and day out. Businesses should focus more on providing a relevant information / suggestion to the users or risk losing them to competitions that provide more focussed and personalized services. Personalization improves customer’s experience, helps drive sales, and increases customer loyalty and retention. Facebook, amazon and twitter are some of the major tech giants that retain and engross its customers with their personalization efforts and it is high time trading firms and brokers catchup this trend.

In trading world, there is an abundance of data, tools and resources available to users. However, the missing piece is personalization. Each trader is different and goes through a life cycle before he figures out what style suits him best. Unfortunately, during this journey many quit early because of initial mistakes. Thus, the aim of this project is to develop an **AI powered data driven personal assistant** to each trader that will analyse his own trading patterns as well as learn from other successful traders to suggest/guide/enable the trader to make better trading decisions.



To create personalized AI signal engine, we combine the user specific data such as historical trades, open positions and behavioural data from the broker and market data such as historical and real time feeds on prices, sentiments, fundamentals, analyst estimates and corporate actions from external data vendor. These data are combined to form a collective intelligence framework which will be used to come up with personalized recommendations/signals/alerts.

**Methodology**: Simple rule-based heuristics, machine learning (supervised and unsupervised) algorithms such as logistic regression, association rule mining, clustering algorithms, CART (Classification and regression trees) algorithms like random forests and highly complex algorithms such as deep neural networks architectures like RNN (Recurrent neural networks) and LSTM (Long short-term memory). The priority however is to use the simplest algorithms where the outcome is explainable to user in plain human language. Complex algorithms will be used only when patters are very highly complex and intertwined.



# **User Profiling**

In order to give most personalised experience, we should really KYC (Know your customer) over and above regulatory KYC. Following is a non-exhaustive list of things we should know about our clients to enable us understand them:

* Who is the customer (scalper/day trader/ swing trader / investor)?
* What does he like to trade (stocks/ forex/commodities/cryptos)?
* Why does he trade (news/technical signals/sentiments)?
* How is he aligned with market (trend following/mean reversion)?
* How much risk is he willing to take?
* Last but not the least – whether he is profitable?

When we know the answers for above, we would be able to provide specific needs of the customers.

## **Profiling a new user:**

In case of a new user, we will not be having much of his trading history to automatically profile him. In this case it would be better to ask the user to fill a very simple questionnaire like below while sign-up:

“Help us to serve you better by answering few quick questions”

1. Which instruments you prefer to trade?

Stocks / Commodities / Forex / Cryptos

1. Which asset classes you prefer to trade?

Spot / Futures / Options

1. Which of the following best describes you?

Scalper / Day trader / Swing trader / Investor

(on mouse-over explain what each kind of trader does)

1. What will be a typical leverage level you use for trades?

1x / 3x / 10x / 20x / 100x

1. What kind of information you trade?

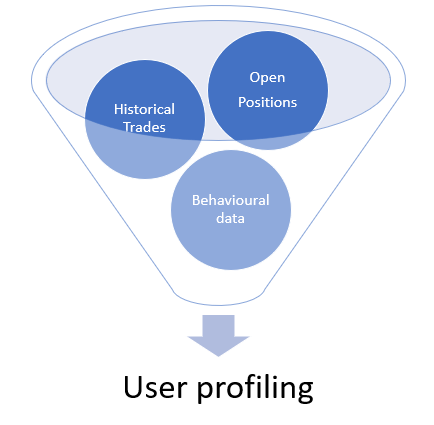
News-Events / Technical indicators / candlestick patterns / Social media trends

(on mouse-over explain what each mean)

While getting these answers, we should also have in mind that the newly joining client will most likely choose multiple options in the above questionnaire (Most curious ones will be open for all) and might even exhibit a completely different behaviour when he starts trading. Hence it is very important to train a bot that will constantly monitor user trades and automatically updates his profile based on trading behaviour.

## **Profiling an existing user:**

For an existing user who made some trades, we collect following data in real-time:



* User trades data:
  + Historical positions:
    - Symbol, position entry time, entry price, Side (Buy/Sell)
    - Exit time, exit price, leverage,
    - If available: List of Target, stoploss
  + Current open positions
    - Symbol, position entry time, entry price, Side (Buy/Sell)
    - Leverage,
    - If available: List of Target, stoploss
* User behavioural data (if available):
  + Does the trader see chart most of the time? If yes following data will be useful:
    - Timeframe most used by trader in chart (1m, 1h, 1D etc.)
    - Chart type most used (Candlestick/ Heikin-Ashi/Line/Renko/Point&Figure)
    - Technical indicators most used (Refer list of technical indicators)
    - Does the trader use trendlines / other chart drawings?
  + Does the trader read more on news etc? If yes what pages inside app he spends most of the time in

Using the above data, we profile an existing user as given in following table. This will be a continuous process as users tend to change their behaviour and more data would lead to more accurate results.

|  |  |  |
| --- | --- | --- |
| **Attribute** | **Types** | **How to classify based on data?** |
| Trading style  (who) | Scalper | Avg. position holding time < 2 hours |
| Day trader | Avg. position holding time < 1 day |
| Swing trader | Avg. position holding time < 1 month |
| Investor | Avg. position holding time > 1 month.  Takes only Long position. |
| Segments  (what) | Stocks | What is the major instrument that the client trades?  Does he have a favourite list of tickers/symbols?  Subclasses:  - Large-mid-small cap stocks only;  Major-minor pairs in fx ; Top-new cryptos etc..  - Futures only/ Options only / Spot only |
| Commodities |
| Forex |
| Cryptos |
| Strategy type  (how) | Trend following | Trades in line with market trend  Ex. Buys above moving average and sells below MA. |
| Mean reversion | Trades against the market trend  Ex. Sells when RSI > 70, Buys when RSI < 70 |
| Trading rationale  (why) | News/Event driven | Trades are circled around key news/event announcements like corporate actions, earnings announcements, economic calendars etc.  - Can also be found if we have data about whether the trader browses more in news and announcement pages |
| Technical indicators | Trades are based on technical indicator signals.  - Can check if the entries are coinciding with technical indicator signals  - Can also know if we have data on whether the trader uses lots of technical indicators while plotting chart |
| Candlestick patterns | Trades are based on support/resistance and other candlestick patterns.  - Can check if the entries are coinciding with candlestick pattern occurrence  - Can also know if we have data on whether the trader uses lots of trendlines while plotting chart |
| Trends/Sentiments | Trades based on trending stocks/stocks in news/ social media trends |
| Risk profile  (how much) | Dynamic (Risk seeking) | Risk seeking trader is the one who is willing to accept large drawdowns for larger returns. Risk averse is opposite where trader cuts short his losses.  We assess the risk profile by following metrics:  - Leverage level  - Avg. drawdown%  - win ratio %  - Standard deviation of trades profits/losses  - Target to Stop loss ratio |
| Balanced (Risk Neutral) |
| Conservative (Risk averse) |
| Performance  (will he stay) | Profitable | Client’s profits/Loss |
| Break-even |
| Loss |

## **Market Conditions**

“History repeats itself” is what most of the traders believe in when they delve into all the past data and base their trades by analysing them. Hence an important aspect of the bot is to constantly monitor the market conditions / events and check how the current scenario or similar events have played out in past. This is critical as identifying and classifying the current market condition / Events will help us in better personalization. Again, how and what we define as an “Event” will completely depend upon what kind of user we are going to recommend them.

Deducing which event has triggered the user to place an order is vital as this directly impacts on what kind of suggestions we are going to present to him. In previous section it was mentioned that behavioural data is optional. If we have behavioural data it would be much easier for the model/heuristics to rightly tag the events that had driven the user to trade. In absence of it, we can still use the trade entry time information and look for specific events in market around the trade entry time and identify the trading rationale.



Following sections provides details on what kind of data that has to be collected/computed and how this data will be used to deduce the rationale behind trading.

## News/Earning announcement/Economic calendar/Fundamentals

* For each of the asset class collect all historical events
  + Stocks: Earnings announcements, Mergers, Stock splits, Major news like CEO change, major project winning, IPOs, FPOs. Fundamentals: P/E ratio, market cap, revenues, cash flow ratios and financial ratios.
  + Commodities: Economic calendars and key number announcement such as production data by major countries
  + Forex: Monitory policy and other money supply/inflation announcements by FED and ECB
  + Cryptos: Forking, Halving, Security breaches, new ICOs
* During these historical events observe how the prices have reflected the information
* In real-time news-feeds look for the above-mentioned events and tag the historical stock reactions to these events.

## Social media sentiments

A section of users might be interested in trading based on social media trends. In fact, there is an article that correlates former president Trump’s tweets with share market returns. This is enhanced now with the reddit groups which pumped the share prices of certain stocks like Gamestop. The idea here is to collect historical social media content relevant to any stock (tweets/reddit group discussions) and check whether some user’s trades are correlating with those stocks. If yes, then we classify those users as sentiment-based traders and in future we could suggest them the stocks that are trending. Messages can be like “DODGECOIN is currently trending in twitter after Elon musk’s recent tweet about it”

## Technical-indicator based events

In historical data and also real time we calculate following technical indicators and generate Buy/Sell signals from them. Each indicator has to be computed for multiple timeframes and multiple parameters and their signals have to stored in history.

|  |  |  |
| --- | --- | --- |
| * Acceleration Bands (ABANDS) * Accumulation/Distribution (AD) * Average Directional Movement (ADX) * Adaptive Moving Average (AMA) * Absolute Price Oscillator (APO) * Aroon (AR) * Aroon Oscillator (ARO) * Average True Range (ATR) * Volume on the Ask (AVOL) * Volume on the Bid and Ask (BAVOL) * Bollinger Band (BBANDS) * Bar Value Area (BVA) * Bid Volume (BVOL) * Band Width (BW) * Commodity Channel Index (CCI) * Chande Momentum Oscillator (CMO) * Double EMA (DEMA) * Directional Movement (DMI) * Exponential (EMA) * Fill Indicator (FILL) * Ichimoku (ICH) * Keltner Channel (KC) * Linear Regression (LR) * Linear Regression Angle (LRA) | * Linear Regression Intercept (LRI) * Linear Regression Slope (LRM) * Moving Average Convergence Divergence (MACD) * Max (MAX) * Money Flow Index (MFI) * Midpoint (MIDPNT) * Midprice (MIDPRI) * Min (MIN) * MinMax (MINMAX) * Momentum (MOM) * Normalized Average True Range (NATR) * On Balance Volume (OBV) * Price Channel (PC) * PLOT (PLT) * Percent Price Oscillator (PPO) * Price Volume Trend (PVT) * Rate of Change (ROC) * Rate of Change (ROC100) * Rate of Change (ROCP) * Rate of Change (ROCR) * Relative Strength Indicator (RSI) * Session Volume (S\_VOL) * Parabolic Sar (SAR) | * Session Cumulative Ask (SAVOL) * Session Cumulative Bid (SBVOL) * Simple Moving Average (SMA) * Standard Deviation (STDDEV) * Stochastic (STOCH) * Stochastic Fast (StochF) * T3 (T3) * Triple Exponential Moving Average (TEMA) * Triangular Moving Average (TRIMA) * Triple Exponential Moving Average Oscillator (TRIX) * Time Series Forecast (TSF) * TT Cumulative Vol Delta (TT CVD) * Ultimate Oscillator (ULTOSC) * Volume At Price (VAP) * Volume (VOLUME) * Volume Delta (Vol ∆) * Volume Weighted Average Price (VWAP) * Williams % R (WillR) * Weighted Moving Average (WMA) * Welles Wilder's Smoothing Average (WWS) |

## Candlestick pattern-based events

In historical data and also real time we look for / compute following candlestick patterns and generate Buy/Sell signals from them. Each pattern has to be computed for multiple timeframes and their signals have to be stored for history.

|  |  |  |
| --- | --- | --- |
| * Two Crows * Three Black Crows * Three Inside Up/Down * Three-Line Strike * Three Outside Up/Down * Three Stars In The South * Three Advancing White Soldiers * Abandoned Baby * Advance Block * Belt-hold * Breakaway * Closing Marubozu * Concealing Baby Swallow * Counterattack * Dark Cloud Cover * Doji * Doji Star * Dragonfly Doji * Engulfing Pattern * Evening Doji Star * Evening Star | * Up/Down-gap side-by-side white lines * Gravestone Doji * Hammer * Hanging Man * Harami Pattern * Harami Cross Pattern * High-Wave Candle * Hikkake Pattern * Modified Hikkake Pattern * Homing Pigeon * Identical Three Crows * In-Neck Pattern * Inverted Hammer * Kicking * Spinning Top * Stalled Pattern * Stick Sandwich * Takuri (Dragonfly Doji with very long lower shadow) * Tasuki Gap * Thrusting Pattern | * Tristar Pattern * Unique 3 River * Upside Gap Two Crows * Upside/Downside Gap Three Methods * Ladder Bottom * Long Legged Doji * Long Line Candle * Marubozu * Matching Low * Mat Hold * Morning Doji Star * Morning Star * On-Neck Pattern * Piercing Pattern * Rickshaw Man * Rising/Falling Three Methods * Separating Lines * Shooting Star * Short Line Candle |

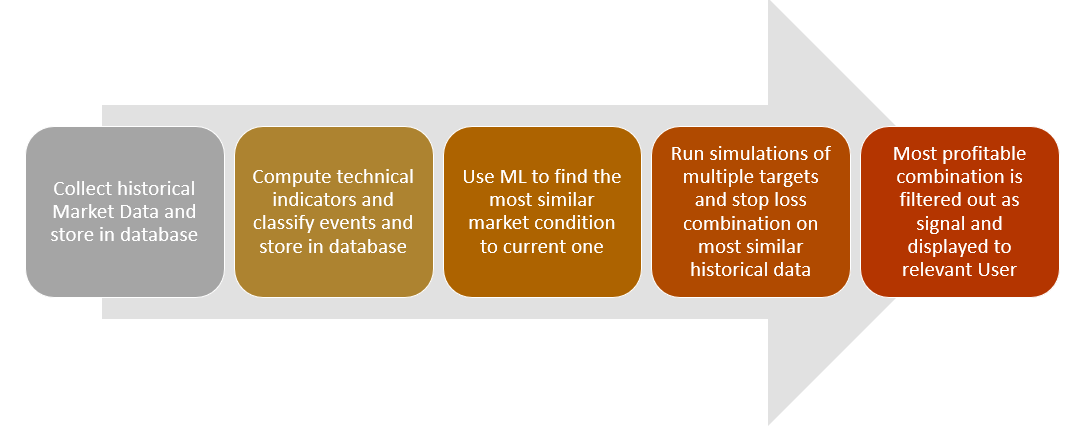
# Roadmap for Implementation

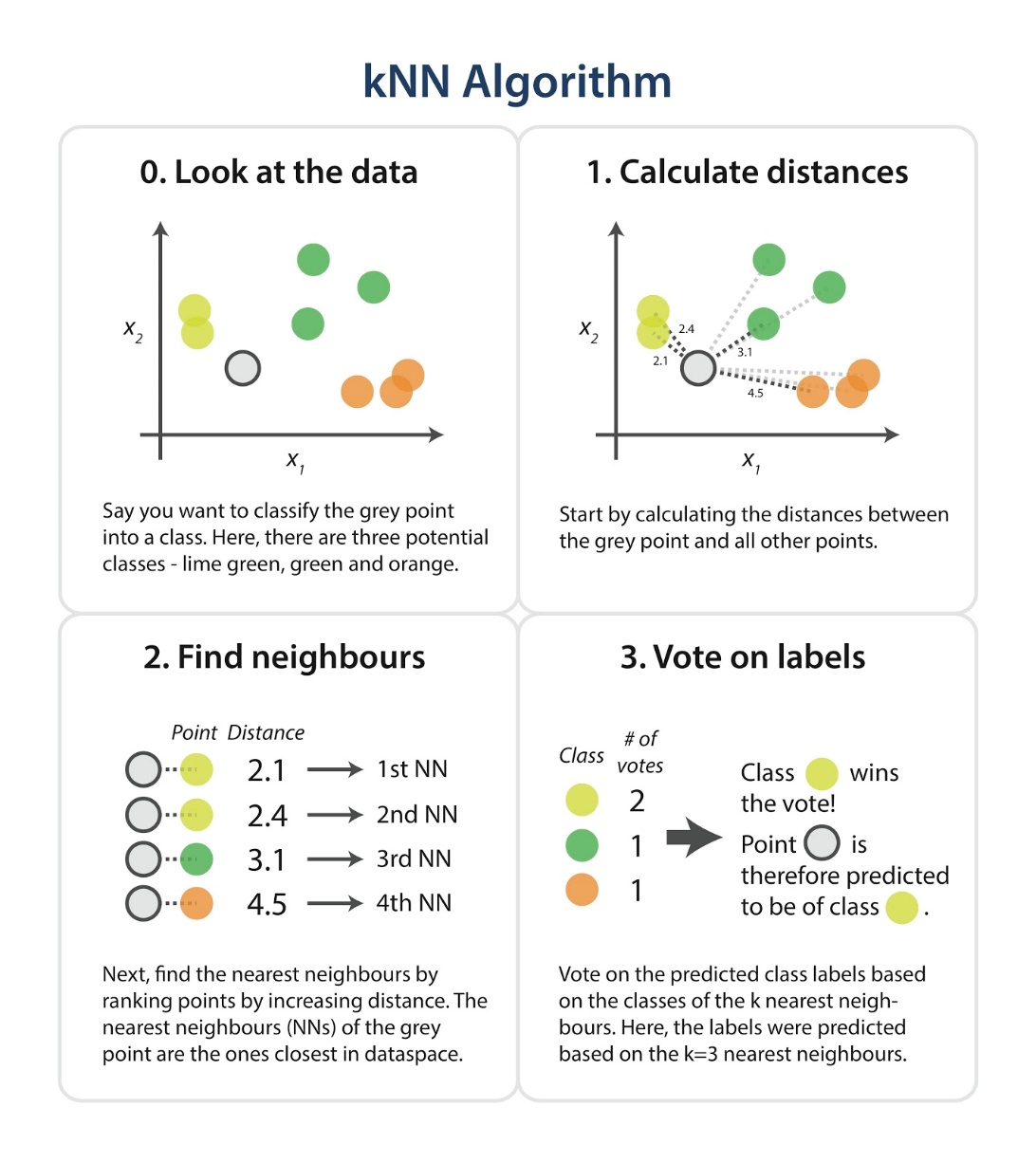
# Stage 1: Project setup and Basic- personalization

When we start fresh, we will not have many users signed up yet. Thus, we will be limited by user specific trades data. All we will have is questionnaire collected from the user about his profiling.

During this stage, priority will be as follows:

* Gather necessary historical market data (Prices/News/Sentiments/Analyst estimates)
  + Setup IEX Cloud subscriptions and data feeds
* Compute technical indicators and candlestick patterns for historical data and upload signals to central data base by datetime.
* Develop a bot that will scan for market conditions similar to current markets using machine learning (like: KNN algorithm). Run simulations of trades on those historical market conditions and shortlist the profitable settings for recommendation.





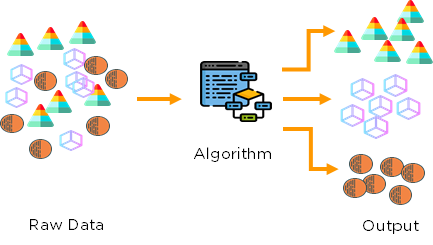
**Example**

* Observe the intraday chart patterns of today’s morning 3 hours of trade in AAPL. Using machine learning KNN algorithm (or similar distance-based algorithms) find most similar days in past.
* Take top 10 similar days and generate/simulate trades with multiple setting of target and stop loss. Find the most optimal setting. (To avoid over fitting, use test/validation sets). Run this for multiple stocks and filter the signals with highest probability of success. Then alert the user with message *“Long AAPL with 5% target and 3% stop loss. This signal has 70% probability of success”*
* Using the data collected from questionnaire we can recommend similar stocks/ industries to the users as well.

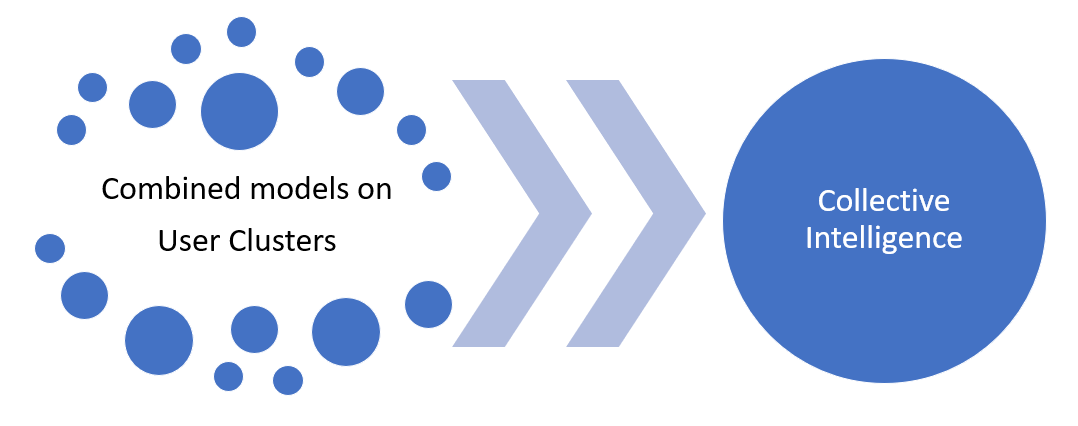
# Stage 2: Collective Intelligence

When we have enough users in system (>100) with considerable trading activity (>1000 trades together), we can start using cross section data of user profiles and market conditions, to put together a framework for collective intelligence. This is achieved by defining user clusters. using rule-based groupings or using machine learning clustering algorithms. In whichever case we cluster, we ensure to have weightage systems as output so that the clusters are interpretable like: Risk seeking forex scalpers; News driven stock investor;   
Sentiment based day traders etc.

**Clustering Algorithm Illustration**



Note: The ultimate aim is to have a cluster with as detailed dimension as possible like – Risk seeking technical indicator-based momentum forex scalpers. However, we could achieve this much detailed cluster only when we have enough number of users within each detailed cluster. This would be possible when we have very high number of users (>5000). At this stage we could see the bot becoming more personalised.



Once we group the users into multiple clusters, these user clusters will aid in advanced personalization. For instance, within each cluster there will be a set of users who would be more profitable and the bot can learn the best practices from these users with respect to when do they enter a trade / which indicator being used / what is their target-stop loss etc. to educate the other loss-making traders.

**Use case: Profitable trading signals**

The bot provides entry signals to the user based on historical analysis of successful trades of multiple users.

**Methodology:**

1. For each cluster of user profile (ex. Risk seeking momentum based crypto day traders), the bot gathers the list of profitable and lossy trades along with market data (events/ indicators/ fundamentals/ sentiments) during entry time of each trade.
2. The bot then uses machine learning to find the pattern behind the profitable trades to come up with a model.
3. The priority is to keep this model as simple and explainable as possible. Example: Whenever RSI > 70 and prices exceed 2.5 Bollinger bands, there is 80% probability that price would fall by 5%. We will go for complex models only when the prediction probability is very weak (<65%) to improve accuracy.
4. This model can be applied in real time market to generate signals whenever a similar market conditions of profitable trades reoccur.

**User message/alert/signal:**

*Short EURUSD with 5% target and 3% stop loss as RSI > 70 and price > 2.5 upper band of Bollinger bands. This trade has 80% probability of success.*

# Stage 3: Complete personalization

Stage 1 and stage 2 are broker specific where total number of customers can drive to complete those stages. Stage 3 is however with respect to a specific user. Here the idea is to analyse the trading behaviour of individual users and personalise their signals and recommendations. This stage can be applied to any user who has completed at least 50 trades. This would enable the system to perform meaningful analysis on trading behaviour of that user. More number of trades would certainly improve performance.

## User-Event mapping

When a specific user is observed to trade the instrument around the certain events (news based/technical indicators etc.), then we can classify the trader as specific event-driven user. This can be done by looking at historical trade entry times of the user and also using machine learning algorithms like “Association Rule Mining”

A same event can be relevant to multiple user profiles. For instance, consider a AAPL if going to announce its earnings tomorrow:

* A message to a news-driven user who has previously have history in trading around news earnings announcements:

*“AAPL” went up by 10% last quarter after announcing a 20% spike in profits. Watchout for this quarter’s announcement tomorrow. Find about analysts estimates here. Find out the open interest here. How other traders feel about tomorrow’s announcements here.*

* On other hand, to a technical indicator-based stock trader who might not be following the news, this news is still important from volatility stand point. Messages to them can be like below:
  + Risk averse trader*: “Amid AAPL earnings announcements, the stock is expected to be very volatile. Please avoid trading in it.”*
  + Risk seeking trader: *“Amid AAPL earnings announcements, the stock is expected to be very volatile and produce many intraday trading opportunities. Happy Trading!”*

Consider another event such as bitcoin prices have moved above 100 day moving average.

* To a momentum-based swing trader message can be like *“Bitcoin prices went above 100-day MA. Buy with 10% target and 5% stop loss. This signal has 65% probability of success in the past”*
* To a men reversion-based intraday trader message can be like *“Bitcoin prices went above 100-day MA by 2%. A short term mean reversion to 100 day MA is possible. Short Bitcoin with 100 MA as target. This signal has 80% probability of success in the past”*

Hence it is important that we specifically mention which segment of event would be relevant to which user and how the alerts/suggestions/signals are altered based on same information but to different user profiles. This has to be done manually involving trading experts.

There can be several personalization use cases built upon using user profiling and collective intelligence.

**Use case 1: Where did I go wrong?**

Identifying loss making customers and suggest them why they made those losses and what they could have done better. By this way we can aid the user to correct his mistakes and retain them.

***Methodology:***

We will identify a trader as consistent loss maker when his average profit is much less than his average loss, and has poor win%.

**User message/alert/signal:**

*“You are cutting short your profits however making your losses to run”* along with a statistic of his historical losses and profits can enable him to reflect upon his mistakes and correct them in subsequent trades. We can also have an article like this <https://www.forextraders.com/forex-education/forex-strategy/trading-strategy-101-cut-losses-short-let-profits-run-on/> for a detailed read out.

**Use case 2: The opportunity is back**

The bot can go through the historical profitable trades of a user and analyse the market if similar entry points are currently available in same or similar instrument.

***Methodology:***

Check the market condition that existed during user’s profitable trade and check if the current market condition is very similar to that one.

**User message/alert/signal:**

*“Remember the trade where you went long did in AAPL on 2nd Oct 2020 and made 5% profit? A similar opportunity is currently found in AAPL, AMZN”*

**Use case 3: Re-consider your stop-loss/target**

The bot can check for open position of the users along with the target and stop loss he has set for this open position to suggest if the stop loss / target is too tight/wider.

**Methodology:**

The bot can scan through similar price and volatility pattern in the market and find the optimal target and stop loss using historical data. This can also be one by using profitable traders’ (from user’s cluster) target and stop loss for similar trades.

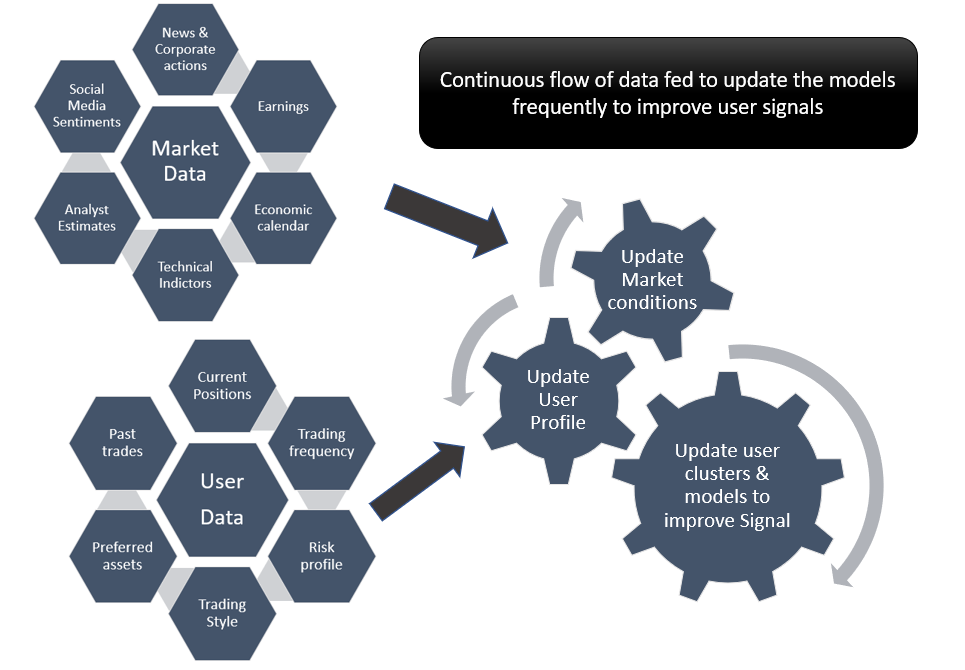
**User message/alert/signal:**

*“The volatility is too high in market and there is 90% probability that current stop loss (1%) will be hit before current target (2%) is reached. Kindly consider a wider stop loss (Recommended stop loss: 3%)”*

Above are just few examples of personalised services that can be provided to users.

# Final working model

The ultimate aim of the project is to deliver a system that continuously feeds user and market data to the bot that updates the user clusters and improves models.



# Running Cost

Options:

Google Cloud: 500$-5000$

Microsoft Azure: 400$-4000$

Hostkey:1500$

# Milestones and Budget

**Milestone 1 - Project set up & basic personalization**

**Time line:**

**Cost:**

|  |  |
| --- | --- |
| **Project set up** | |
| Setting up project cloud workspace | 1 week |
| Setting up databases | 1 week |
| Finalise vendors for market data – stocks / forex / cryptos / commodities: Both historical and live feeds | 1 week |
| Finalise vendors for fundamentals, news, sentiments, analyst estimates |
| Data warehousing and engineering to set up all data in cloud | 1 week |
| **Backend development- Data Ingestion and featurization** | |
| Write code to data ingestion, cleaning and transformation | 1 week |
| Code to sample data to multiple time frames | 1 week |
| Code to technical indicators, candlestick patterns | 1 week |
| **Backend development – Core algorithm** | |
| Develop ML algorithm to find similar market condition in past | 2 weeks |
| Simulate historical trades using technical indicators, news events etc. to filter out all successful combinations of target-stop loss for scalper, day trader, swing trader and investor | 2 weeks |
| **Deployment** | |
| User profiling questionnaire and create test users | 1 week |
| Using algorithm, develop live signal generator bot with live feeds |
| Customise user specific messages |
| Deploy for test users and test the product |
| **Deliver MVP** | |

**Milestone 2 - Collective Intelligence**

**Time line:**

**Cost:**

|  |  |
| --- | --- |
| **User Profiling and clustering** | |
| Tagging a user’s trading frequency | 2 weeks |
| Tagging a user’s risk level |
| Tagging a user’s asset and trading preferences |
| Tagging behavioural aspects of user |
| Tagging user’s strategy style |
| Deducing if the user’s trades rational using ML | 3 weeks |
| User clustering based on trading rational and profiles | 2 weeks |
| **Collective Intelligence** | |
| AI algorithm to mine through profitable trades of users and come up with an intelligent signal generator | 3 weeks |
| Mapping the signal generator to user profiles to show relevant and personalized signals | 1 week |
| Live signal generator and user message sender | 1 week |

**Milestone 3 – Complete personalization**

**Time line:**

**Cost:**

|  |  |
| --- | --- |
| Come up with personalization use cases (See stage 3 for examples) and setup infrastructure | 1 week |
| Algorithm to develop the use case | 2 weeks / use case |
| Deploy the use case and go live | 2 weeks / use case |