IIMA
विद्याविनियोगाद्विकास:

Brij Disa Centre for Data Science & Artificial Intelligence

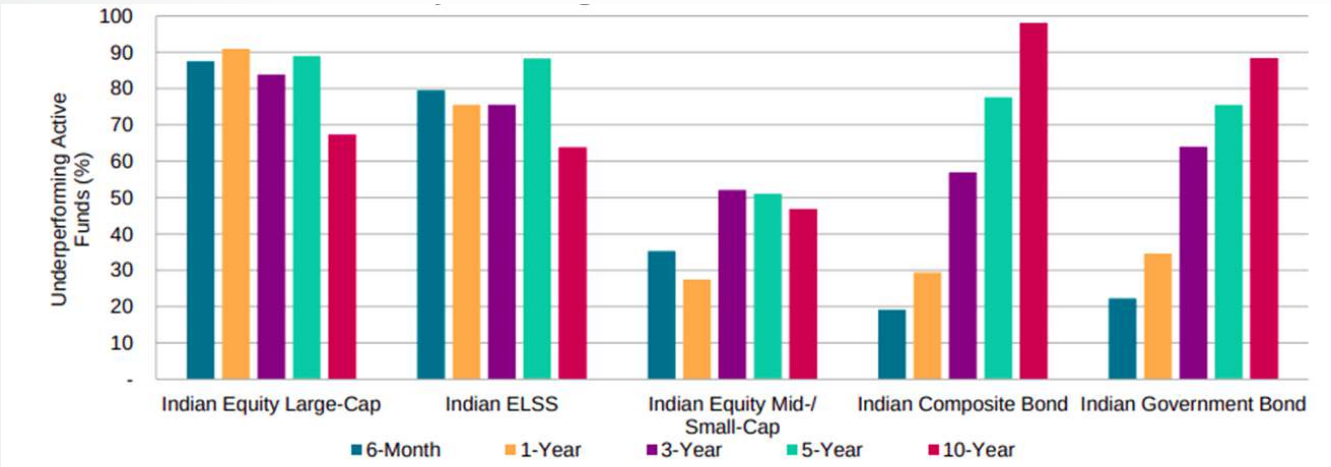
Communique
January, 2023

CREATING A PORTFOLIO USING A MEAN-REVERTING PAIR TRADING STRATEGY

Vaishnav Garg

Vaishnav Garg is a Pre-doctoral researcher at IIM Ahmedabad working with Prof. Anindya Chakrabarti. He holds a Bachelors and Masters degree in Mathematics and his research interests lie in fields of application of quantitative methods in Finance such as Financial Econometrics and Mathematical Finance. Other research interests include network analysis of Financial market, asset return properties and volatility modelling.

Portfolio management, simply put, is the process of selecting the right investment opportunities for an individual or any concerned entity with the aim of maximizing the returns while keeping the risks as low as possible. This task, however, is not very easy and requires the expertise of the most highly qualified "fund managers," who are paid hefty sums for the job. However, if we look at the performances of the Active funds in India, a surprising majority of them fail to surpass their benchmark indices. According to the reports published by SPIVA (S&P Indices Versus Active funds) India in July 2022, about 70% of Large-cap equity funds underperform the BSE100 index over a 10-year investment period, and the index outperforms about 90% of these funds over a shorter time horizon of about six months-1 year. This inability of funds to outperform their benchmarks is not limited to India but can also be seen for Funds worldwide.



Source: S&P Dow Jones Indices LLC, Morningstar and Association of Mutual Funds in India. Data as of June 30, 2022. Outperformance is based on equal-weighted fund counts. Index performance based on total return in INR. Past performance is no guarantee of future results. Chart is provided for illustrative purposes.

Based on the goals and specifications of the stakeholders, various categories of portfolios can be created, which may attract investments across multiple asset classes such as Equity, Mutual Funds, ETFs, Bonds, Cryptocurrency, Gold, Real Estate, etc. However, this article focuses on creating only an equity portfolio. Here, we will demonstrate a portfolio that uses a simple yet powerful statistical measure that outperforms its benchmark index throughout the period of study.

Market Players

Before diving in, let's first learn about the market players involved in trading assets and the various strategies often used by them. Since our focus is on the equity market, there are two trading strategies deployed by these traders – Fundamental Analysis and Technical Analysis. At any given point, there are 2 groups of traders called Fundamentalist and Chartist, respectively, based on their trading strategies. The fundamentalists look at the company's strengths and aim to determine the intrinsic value by pricing them using parameters such as revenue, earnings, future growth, profit margins, etc. The chartists rely more on the patterns of stock movement and volume traded to forecast the direction of price movement. A chartist views a stock to be in either of two regimes – a mean-reverting or a momentum phase. A momentum regime assumes that the stock prices will constantly rise or fall, whereas a mean-reverting regime assumes that the stock will revert to its long-term equilibrium after a sudden rise or fall. In a momentum regime, a trader will take a long position when prices increase and take short positions when prices decrease, whereas, in a mean-reverting regime, a trader will take a long position when stock prices fall and will short when stock prices increase.

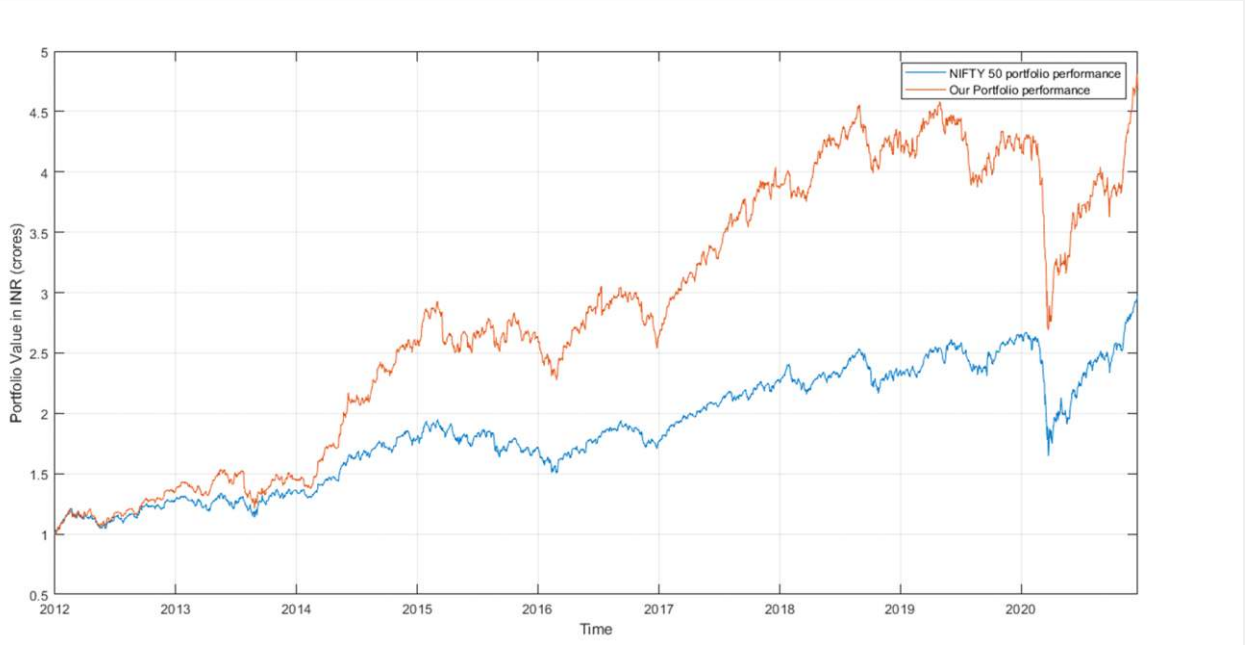
Pair Trading

Pair trading is a trading strategy requiring a trader to buy a particular asset, short another related asset simultaneously, and try to take advantage of statistical arbitrage due to their price differentials. It is believed that the first pair trading was done by Morgan Stanley in the early 80s. For example, a trader following a correlation-based Pair trading strategy (a type of pair trading strategy) would select two similar stocks, say Coca-Cola and Pepsi. These two companies are of similar size and presence in India with a similar utility and client base. We can safely say that whatever business environment affects Coca-Cola will affect Pepsi and vice-versa. Hence, we will observe that their stock price movements are similar, and on a given day, an x% increase in Coca-Cola stock price will accompany a y% increase in Pepsi stock price. Let's assume on a particular day, Coca-Cola stock prices increase by x%, but Pepsi stock price doesn't increase. This makes the Pepsi stock intrinsically cheap and Coca-cola stock inherently costly. So a pair trader will buy the Pepsi stock and short sell the Coca-cola stock expecting the stocks to come in equilibrium and profit on this supposed abnormality. The trader exploits the price differentials between two correlated stocks using the correlation pair strategy.

Constructing our Portfolio

We use an amalgamation of the above ideas to create a portfolio with a mean-reverting pairs trading strategy, i.e., we are trading as a chartist using a pair trading strategy hoping to make profits from the statistical arbitrage of mean-reverting property of stocks. This is purely a historical data-driven trading strategy and does not rely on any fundamental analysis of the companies. Since we focus only on large-cap stocks, we will form a portfolio using the NIFTY50 companies. We exploit the mean-reversion property of pair of stocks where we take a long position on one stock and short on the second stock such that the resultant has the highest possible mean-reversion tendencies. To measure the mean-reversion tendency, we use the Hurst exponent, which indicates whether a stock is in its momentum or mean-reverting regime. Since we have 50 stocks in NIFTY, there are 1225 pairs possible, and we choose the best 10 pairs out of them to invest.

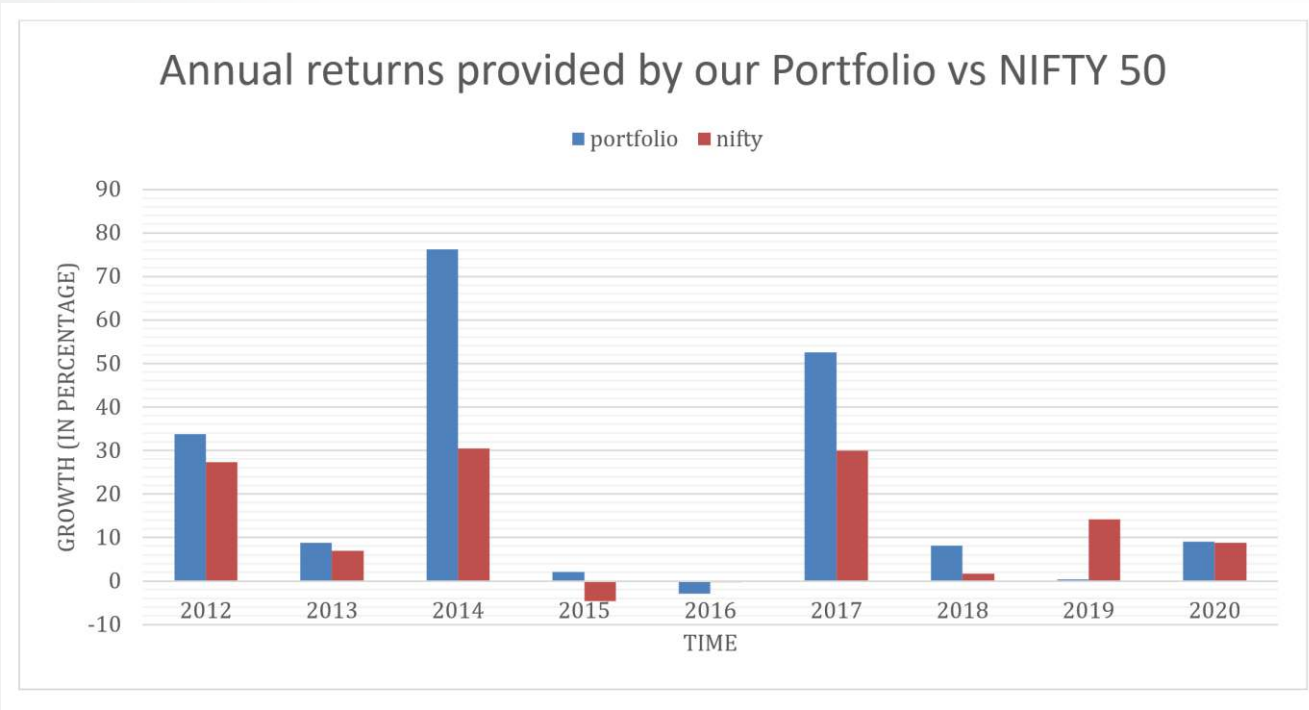
We construct our portfolio at the start of each calendar year based on the performance of stocks in the previous year, i.e., the 10 pairs we selected to invest in 2012 was determined by the performance of NIFTY50 stocks in 2011. We invested a sum of 1 crore in the portfolio, equally distributed in 10 pairs at the start of January 2012, and shuffled the portfolio at the end of each calendar year. We track the performance of this portfolio throughout the year. At the end of each year, since we are recycling our whole portfolio, we levy a hefty 2% transaction fee on the current portfolio value (transaction costs for buying and selling assets). We compare our portfolio to the benchmark index in the next section.



Comparison

We consider our benchmark index, NIFTY 50, and compare its returns to our portfolio returns, assuming we had invested the same sum of Rs. 1 crore in an index fund that mimics NIFTY 50 performance. Had we invested the same amount in 2012 in our portfolio and the index portfolio, the value of our portfolio would always have been greater than the index portfolio. We present the plot of our portfolio against the index above to compare their performance over the time horizon of 2012-2020. An initial investment of 1 crore INR in our portfolio in January 2012 would have compounded to 4.67 crores INR at an average of 18.68% annual growth. In contrast, the same investment of 1 crore INR in an index portfolio would have grown to 2.97 crores INR at an average of 12.90% annual growth. Hence, a passive buy-and-hold investor investing in 2012 in our portfolio would earn an average of 18.68% CAGR growth compared to a growth of 12.90% CAGR in the index portfolio, an enormous difference of 5.78% of annual compounding growth.

We also compare the yearly performance of our portfolio by calculating the annual returns of both portfolios during that period. We present a bar graph below that compares the annual growth in the value of the investment. We observe that our portfolio beat the benchmark every year in our period of study, except in 2019 and just by a margin in 2016. Our portfolio, on average, gives an additional 8.17 % of returns compared to NIFTY50. The analysis also shows that if you are a long-term passive investor who practices a buy-and-hold strategy, our portfolio strategy will significantly outperform the benchmark regardless of the time, date, or year you invest.



References

1. Quynh Bui, Robert. (2022). Applying Hurst Exponent in pair trading strategies on NASDAQ100 index
2. Sergey, Drozdov. (2017). Fractal Optimization as Generator of Market Neutral Long-Short Portfolio
3. Requena, Segovia, Granero. (2017). Introducing Hurst exponent in pair trading
4. Requena, Segovia, Granero. (2020). Some Notes on the Formation of a Pair in Pairs Trading
5. "Pair Trading Logic" by Zerodha from https://zerodha.com/varsity/chapter/pair-trading-basics/
6. "SPIVA India scorecard 2022 mid-year update on performance of active mutual funds vs their benchmarks" from https://tradingqna.com/t/spiva-india-scorecard-2022-mid-year-update-on-performance-of-active-mutual-funds-vs-their-benchmarks/1377

EMERGING CONCERNS WITH SOCIAL MEDIA USAGE

Dr. Kulvinder Kaur

Dr. Kulvinder Kaur is a Post-Doctoral Research Associate at the Indian Institute of Management, Ahmedabad (IIMA). She is working under the Brij Disa Center for Data Science and Artificial Intelligence. Her research interest is to explore the use of social media in varied contexts. She is working on various qualitative and quantitative research projects, and her work has been published in reputed international journals including Journal of Business & Industrial Marketing, The TQM Journal, Benchmarking: An International Journal, etc. She has attended various national and international level conferences and workshops.

Social media have become a critical element of the work and lives of billions of people worldwide. Social media platforms enable paradigm shifts in building relationships, interacting, collaborating, trading, and sharing information. It allows people to communicate with a worldwide audience and has become a major source for real-time information sharing. However, there are numerous negative consequences associated with social media use, including privacy concerns, distraction, addiction, mood swings, sleeping disorders, and technostress, to name a few. Some of the emerging issues associated with social media use are discussed below:

- Researchers are using the term "catfishing" to highlight the issue of the creation of a fictitious online identity for the purpose of deceiving someone [1]. Catfishing is a sort of deceitful behaviour in which a person creates a false identity on a social media account, often to target a specific victim for abuse, deception, fraud, or other purposes. People are often made to believe that they are conversing with real identities, but in reality, they are conversing with a catfish, a person who has stolen the online identity of another person and is exploiting it for his or her own benefits.
- These days, people have access to the same level of telecommunications resources that only major corporations could afford a few decades ago. The ever-increasing access to technology has resulted in its inappropriate usage by harming, shaming, embarrassing, and abusing others. This phenomenon is known as "cyberbullying," and it can be far more dangerous than traditional bullying because it uses a diverse range of media such as manipulated content, images, video, and so on. Cyberbullying is deliberate and repeated online bullying through the use of electronic means, digital and social media platforms [2]. Although it takes place in cyberspace, its consequences could be seen in the real world, including unpleasant feelings such as despair, anger, frustration, shame, and fear. In addition, cyberbullying has been associated with low self-esteem, suicide thoughts, aggressive behaviour, drug abuse, and criminal ideations. Cyberbullying is sometimes interchangeably used with "cyberstalking," which is a form of cyberbullying that violates a person's right to privacy. Cyberstalks generally employ internet databases, social media, and other online resources to track, harass, and threaten their victims.
- Doxing is the online disclosure of personal information about an individual or an organization. This information is usually taken from public databases or online profiles, but enemies may also hack social media accounts to obtain such data. Doxing is a kind of online harassment that involves revealing someone's true name, address, occupation, or other identifying information to the public. Doxing occurs without the victim's permission, with the intent to humiliate or threaten them.
- Many countries are using social media as a surveillance tool. A recent study shows that thirteen of the fifteen Asian nations are either using or developing social media surveillance systems [3]. Although social media sites are in the public domain, using them for surveillance raises ethical concerns regarding privacy protection. Illegitimate surveillance and surveillance abuse are the major concerns for society. It involves the illegal monitoring of an individual's activities, which violates the social norms of society. Special laws should be enacted to safeguard human rights and safety in the era of artificial intelligence monitoring.

- The digital divide is a pervasive and intricate phenomenon in our society. It includes digital inequalities regarding accessibility, usage, and consequences. There is a substantial disparity between those who have access to technology and those who do not. Despite improvements in internet availability in India, the digital divide between men and women remains enormous, as just 33% of women in India use the internet, which is much fewer than males (57%). This disparity is significantly greater between rural males and women (49% versus 25%) [4]. This existing gap has also resulted in inequality and underrepresentation of certain groups on social media platforms, including women, minorities, the elderly, and rural populations.
- Digital vigilantism is a phenomenon in which individuals are collectively aggrieved by the actions of other citizens and organise vengeance via mobile and social media platforms. The offences range from minor social norm violations to terrorism and participation in riots and criminal activities. Social media vigilantism is often judged by the online community to have violated standards and may have real-world repercussions.
- Misinformation is not a new phenomenon, but the emergence of social media has facilitated the quick dissemination of misleading and inaccurate information. Due to the low barriers to information generation and transmission, these platforms provide an ideal ground for spreading misinformation and disinformation, resulting in detrimental consequences for society and posing significant challenges for platform providers and users. Almost every country is facing this issue, which has awakened governments and other organisations worldwide to tackle it. According to a recent Microsoft report, India continues to be the country with the highest number of fake news instances worldwide [5].

Mis-and Disinformation: Key difference and kinds

Misinformation refers to the unintended creation or distribution of information, whereas disinformation is the deliberate fabrication or diffusion of content. The primary distinction between the two is the presence of an intention to deceive. Misinformation and disinformation are used interchangeably for related terms, like conspiracy, fake news, propaganda, and mal-information [6]. While all these terms have a few similarities, the purpose for which the content is created or shared differs, ranging from entertainment purposes to financial gain or to harm someone. A brief description of kinds of mis-and-disinformation is given below (see Figure 1.1).

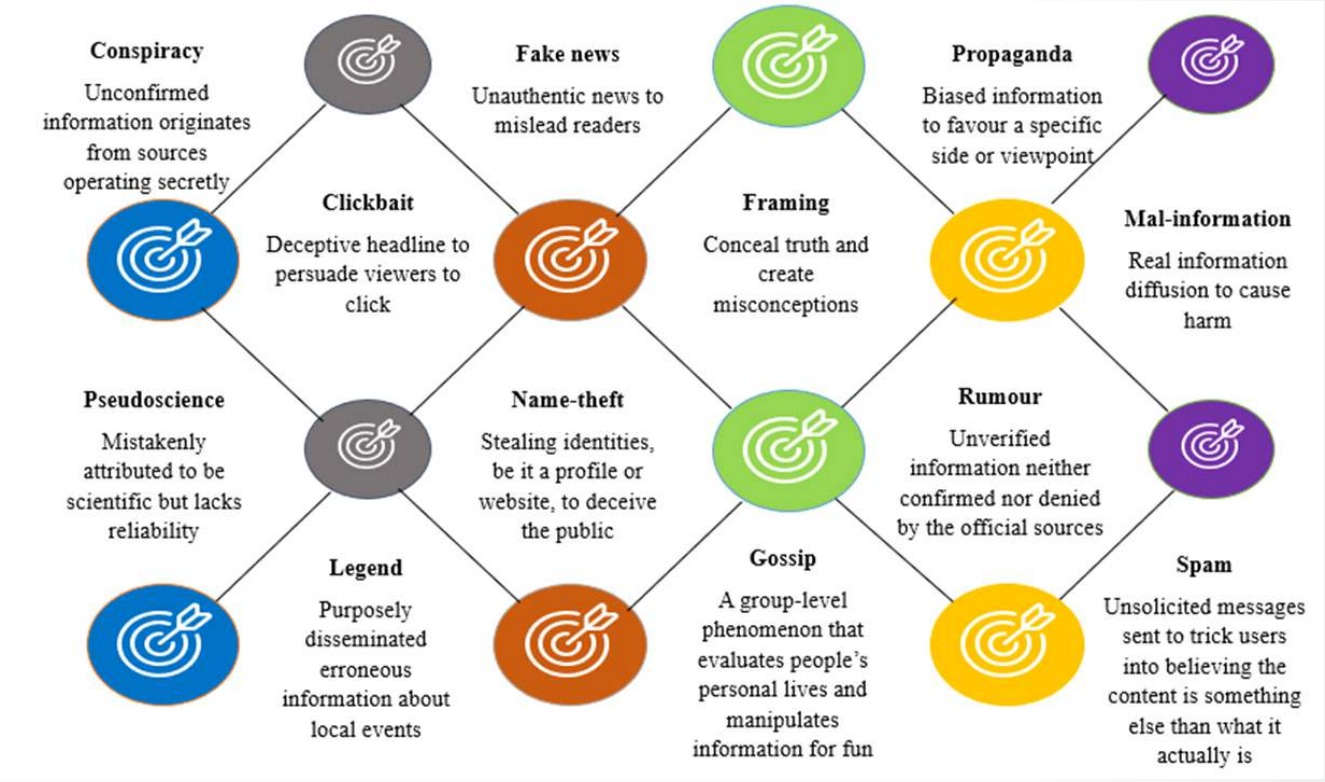


Figure 1: Kinds of mis-and-disinformation

Diffusion, detection, and combating misinformation

Social media users are increasingly exposed to unverified content that is difficult to validate. The content can take various forms, including false text, images, and videos, and can be created in many ways, including by humans, machines, or a combination of the two. There is a rising worry about the propagation of misinformation on social media. Whether intentionally or ignorantly, a massive amount of digital misinformation is disseminated by social media users. The use of logical phrases such as "an expert claimed" and "based on research" persuades them to adopt biased or inaccurate beliefs. Altruism is one of the top motives for the diffusion of information on social media. Information sharing, sociability, information searching, and passing time are other reasons for content sharing. It is essential to double-check the information and consider its accuracy before believing it or forwarding it to other members of one's network. Social media fatigue, laziness, a lack of social media experience, and information overload create hurdles in the authentication process. Early identification of misinformation is crucial to reduce its influence and mitigate the harm caused by such information. Therefore, efforts should be made to spread accurate information as quickly as possible to nullify the effect of fake information.

One of the most severe challenges for social media companies today is figuring out an appropriate detection mechanism. The misinformation detection process is usually dynamic in nature. The detection models are divided into four categories: content-based (text, image, and video), context-based (time and place), propagation-based (circulation modes), and early detection (at the beginning). Most detection systems presuppose some model, which complicates the detection process, so a training protocol is necessary. Using AI and machine learning techniques can help the platforms identify mis-and-disinformation. Other methods for counterfeit information detection include automatic fact-checking through Natural Language Processing or under expert guidance to assess the content's credibility. It is worth mentioning that one's political orientation drives credibility and sharing biases, and conservatives' credibility bias boosted their sharing prejudice. Reliability and credibility are crucial elements in the context of trust in information, which further promote information sharing depending upon the factors influencing users' intentions to share. Users should understand social media interfaces and develop verification behaviours before relying on or sharing content on their digital network. It is essential to identify those who are creating and circulating misleading content. It is also vital to assess fake social media profiles and immediately take strict action. Depending on the detection mechanism, one should carefully choose the right tool to detect fake information.

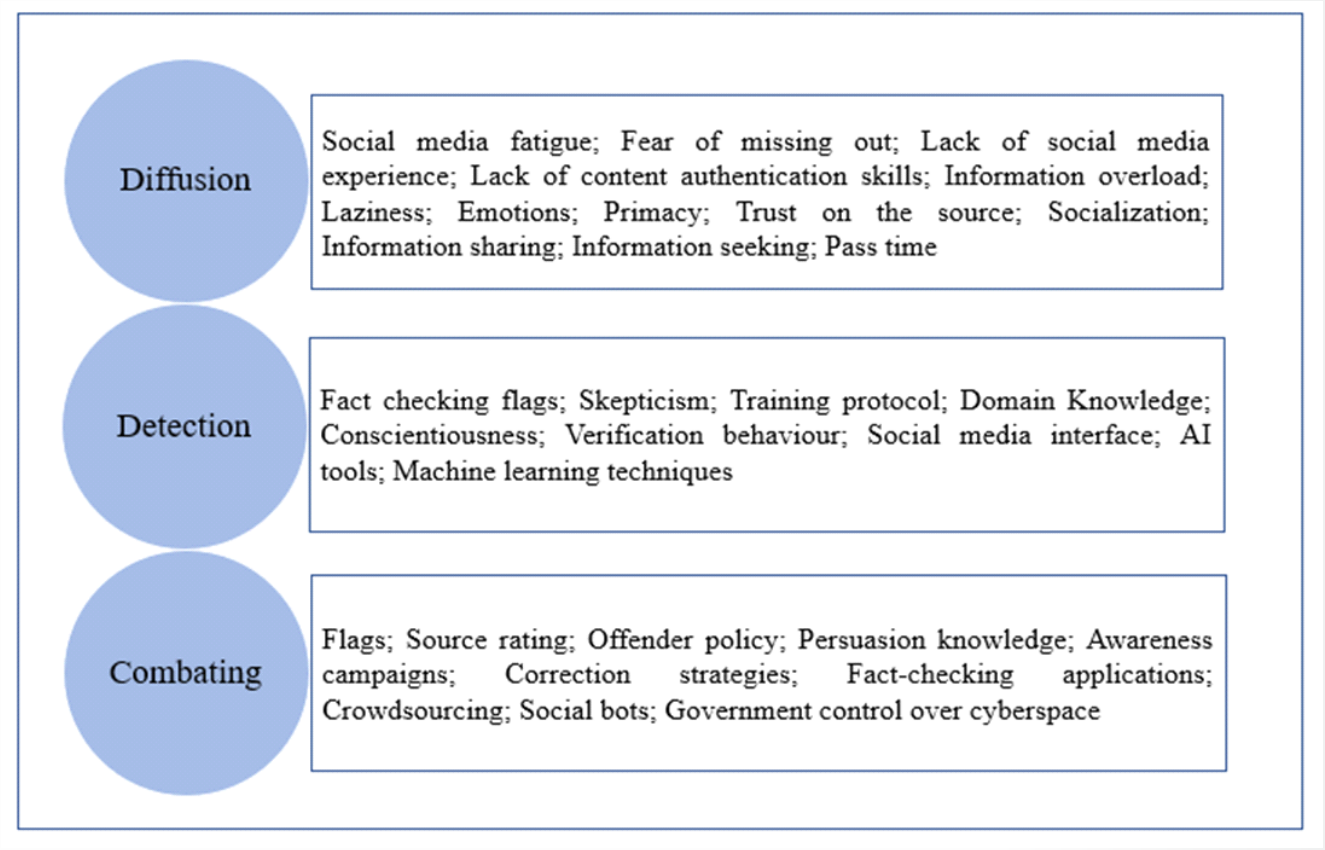


Figure 2: Diffusion, detection and combating misinformation

Researchers, policymakers, and many organisations must take strict actions to minimise the proliferation of misinformation. Source rating, social bots, flagging, and fact-checking applications are a few ways to counterfeit misinformation. Social media platforms, including Facebook, Instagram, LinkedIn, and Twitter, are also creating and modifying their policies and strategies to recognise spam messages, fake profiles, and inappropriate and abusive content. The negative effects of social media necessitate careful regulation of these platforms without jeopardising society's privacy and safety. To combat misinformation, cyber cells, government officials, researchers, policymakers, and international organisations must work together.

References

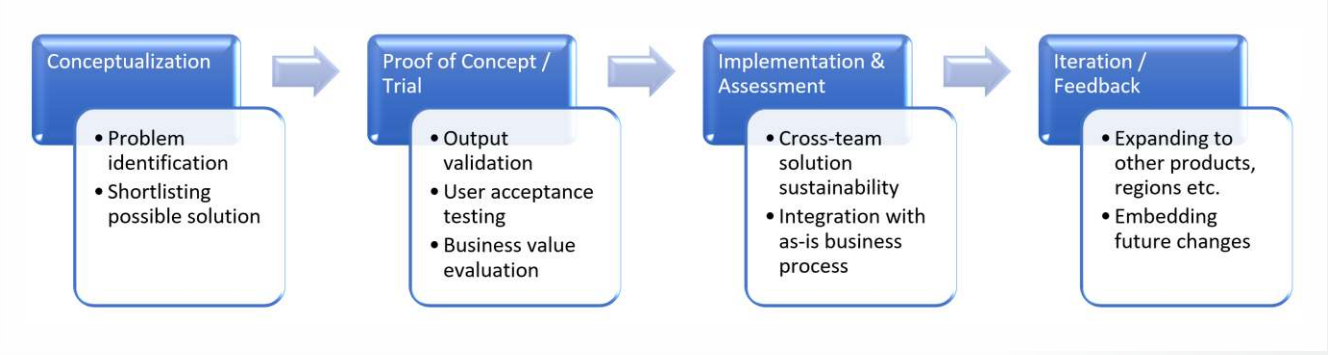
1. Knafo, D. (2021). Digital Desire and the Cyber Imposter: A Psychoanalytic Reflection on Catfishing. Psychoanalytic Dialogues, 31(6), 728–743. https://doi.org/10.1080/10481885.2021.1976187
2. Qureshi, I., Bhatt, B., Gupta, S., & Tiwari, A. A. (2022). Introduction to the Role of Information and Communication Technologies in Polarization. In Causes and Symptoms of Socio-Cultural Polarization (pp. 1–23). Springer Singapore. https://doi.org/10.1007/978-981-16-5268-4_1
3. Shahbaz, A., & Funk, A. (2019). Social media surveillance, retrieved from https://freedomhouse.org/report/freedom-on-the-net/2019/the-crisis-of-social-media/social-media-surveillance
4. McDougal, L., Raj, A., & Singh A. (2022). The digital divide and is it holding back women in India? Gender Project, retrieved from https://www.hindustantimes.com/ht-insight/gender-equality/the-digital-divide-and-is-it-holding-back-women-in-india-101641971745195.html
5. “India has more fake news than any other country in the world: survey”, retrieved from https://www.businessinsider.in/india-has-more-fake-news-than-any-other-country-in-the-world-survey/articleshow/67868418.cms
6. Kaur, K., & Gupta, S. (2022). Towards dissemination, detection and combating misinformation on social media: a literature review. Journal of Business & Industrial Marketing. https://doi.org/10.1108/JBIM-02-2022-0066

CHALLENGES DURING THE IMPLEMENTATION OF AN OPERATION RESEARCH (OR) PROJECT

Abhik

Abhik is currently leading the Optimization & Simulation team in the Advanced Analytics Office at Western Digital, India. He has previously worked with Bayer, FICO, Intel where he applied mathematical optimization to solve challenging problems in different industries.

Various challenges lie in the lifecycle of an **Operation Research** (OR) project . This article focusses mainly on the difficulties encountered during the implementation phase. Often a successful OR model is either not implemented, or the implementation process becomes an arduous journey. Implementation of an OR model is not a one-time exercise, rather it is a continuous process and involves significant hand-holding between the OR modeler, business team and IT team specially during the initial period of implementation. Continued patronage and usage of the OR model depends not on the mathematical novelty of the model, but rather on the quantum of business value generation and the perceived user-friendliness of the OR application in the eyes of the user.



Key Stakeholders	Business Team IT Team	Business Team	Senior Data Scientist IT Team	Business Team
Key Challenges	Awareness about OR	Solution not per user expectation	Integration challenges	Adoption challenges
	Resistance to change		Steep learning curve in the beginning	Investment of time, effort, money in future to make model changes
			Direct revenue impact	Model output acceptable to end user, but not to other stakeholders
			Team/Individual benefits prioritized over organization goals	

Some challenges faced during implementation phase are unique to OR projects. Here we try to highlight some of these challenges.

Conceptualization

Lack of awareness

Because of its mathematical content, OR projects run the risk of not being fully understood by a user without a strong quantitative background. As OR is a relatively new field and still at a very nascent stage, most users are probably seeing a solution of this form for the first time. The chance of adoption of an OR model is higher if the user has experienced the implementation and success of an OR project earlier.

Users' resistance to change

This could stem from multiple factors:

- The user might have been following the manual process for years and are complacent that their current operational process has no defect.
- Thinking that an outsider cannot learn their process and even improve on it.
- Often the business process consists of lot of complex rules and users are not confident if the model can imbibe all those rules.
- Implementation of the OR model would mean more work on their plates – they would need to learn and operate a new software!

Proof of Concept / Trial

Solution not per user expectation due to different reasons

Poor synchronization across teams in the organization, or gap in understanding with the OR team which could lead to the following. Sometimes members of different teams, or even members of the same team are not aligned in terms of requirements or expectations from the model. This may lead to:

- All requirements (objectives, constraints) not clearly communicated
- Problem scope not clearly defined
- The model uses faulty input data

User's intuition that optimizer solution is worse than the manual solution, this could be due to the following:

- While comparing manually derived output with the optimization output, users often look at 'local' solutions and decide what is best.

  Example: In a production planning problem where we try to meet demand of different items, for a particular set of items, the optimizer solution might be worse than the manual solution. While the user may not be satisfied with this outcome, what needs to be compared is the global solution i.e., the solution across all the products instead of looking at a smaller subset of the problem.
- Some constraints are coded as hard constraints, which the manual solution may have considered the constraint as soft.

Simplifications / assumptions to avoid mathematical complexity or reduce solution time:

- o Deterministic approximation of stochastic problem
- o Linear approximation of non-linear problem
- o Stopping criteria to obtain solution faster which may lead to sub-optimal solution

The model ignores 'practical' constraints – these are constraints which are not explicitly stated by the business, but must be included for an operationally feasible solution.

- Allocation of very small or large values

  Example: For a machine-job assignment problem, the model might suggest allocating a job to a machine for a very small time (say 1 minute). In this case, we can add a constraint that if a job is assigned to a machine, it must be for at least 30 minutes (i.e. either 0, or >= 30 minutes).
- The model does not consider 'distribution' constraints. As OR models care only about maximizing or minimizing the objective function, it might allocate an unreasonably high value to a variable with high objective function coefficient, and 0 to variables with low coefficients. Example: The model might suggest selling all the inventory of 'Type A' to a particular customer because this customer pays the highest, but it might not be a good solution because all the customers might want some portion of 'Type A' as otherwise they would move to other competitors. Although the model behaves perfectly in this case, some constraints are required to ensure fair share distribution across customers.

Implementation and Initial Assessment

Too many changes required in the current system for integration with the OR model

The changes could be in the IT architecture or business process.

Users need to invest significant time and effort in the initial stages

During the initial phase of deployment, users need to spend significant time and effort to understand how the model behaves. Often input parameters and objective function coefficients need to be modified by the user for the best output. Different users might require different parameter values, so a one-size-fits-all value would not work. Which value works best is not super-intuitive and can be best judged only after several rounds of experimentation.

Users need to consider the fact that this additional time spent in the initial stages would be offset by the enormous time savings once the model is fully operational.

Team/Individual benefits prioritized over organization goals

Such products are often planned at a higher level or by a separate team working in silos. They look for a solution which might be optimal for individual teams but would be sub-optimal when looking at the overall organization. Resources (example: factory production capacity, inventory to be sold) are limited, hence teams must collaborate to make the utilization of the resource most effective.

Direct revenue impact

Some OR models are very tactical in nature and directly impact the company's revenue (example: pricing models and revenue management models). In such cases, users are hesitant to implement the model output as any negative outcome of the model would be irreversible.

Post Implementation / Iteration Loop (Human and Automated)

Adoption Challenges: Ease of use of the OR application

OR models which are run by users often have a front end. It might happen that the model satisfies all the client requirements, but the front-end is too cumbersome.

Investment of time, effort and money in future to make model changes

While the OR model may be handling all aspects of today's problem, users are unsure whether it can sustain future changes. If the changes are significant, it would imply that lot of time, effort and money would be again required in future.

Model output acceptable to end user, but not to other stakeholders

The user may be content with the solution but are not very convinced if other stakeholders would accept and implement the solution.

Example: In an inventory optimization problem, the inventory planner might be content with the output, but is worried that the procurement team may not be willing to purchase all the raw material allocated by the model

RESEARCH PROJECTS

New Projects:

Financial networks from big data: A multivariate time series based approach

Prof. Anindya Chakrabarti

Financial markets exhibit non-trivial comovement and dependency structure. The standard approach in the finance literature is to consider the market in its aggregate form. A more recent 'data'-oriented approach emphasizes a more granular decomposition of the market so that the aggregate dynamics can be broken down into contributions arising from individual assets. This leads to two analytical problems. First, one has to necessarily deal with a large amount of data such that the process scales with the volume of data (large N and large T where T>>N). Two, analyzing such a large volume of data requires toolkits which are at the intersection of econometrics and machine learning. In this project, the goal is to construct large scale financial networks based on multivariate time series data to capture the dynamics of the system. The main idea is to provide an algorithmic approach to convert time series into networks such that the properties of time series are also inherited by the resulting network. The spectral structure of the comovement network is known to capture, at least partially, the booms and busts in the markets. Here, we take up two specific problems. One, how reliably does the spectral structure reflect the system for the case where T~N. Two, a large chunk of the literature on networks construction depends on bivariate modelling which is subject to failure due to multiple hypothesis testing. Therefore, an imminent question is how to construct a network with a direct multivariate model.

Multi-period Facility Interdiction Problem

Prof. Sachin Jayaswal

We propose to study a multiperiod interdiction problem, in which the leader (attacker) with a limited interdiction budget decides the sequence of facilities to interdict (destroy) over time so as to inflict the maximum cumulative damage to the follower. The follower's objective is to serve a given set of demand points from the surviving subset of facilities the minimum cumulative cost across all periods. for this, his decisions include the assignments of demand nodes to the surviving facilities and the allocation of his limited budget to the revival of interdicted facilities and the protection of the surviving facilities against their interdiction in the future periods. The multi-period version of the problem, which is the focus of the proposed study, presents additional complexity due to the leader's interdiction decisions constrained by the follower's protection decisions. The objective of the proposed study is to design efficient exact solution methods for this challenging bilevel integer program.

Data-driven auction design: A computational approach

Prof. Jeevant Rampal

Details:Auctions are often used to sell property rights for liquor licenses, spectrum licenses, land and mineral rights, and construction projects etc. This project investigates potential improvements in these auctions using a computational data-driven approach. The first part of this project will be to collect primary data of the participants and their choices in auctions. Subsequently, using the game-theoretic properties of the chosen auction design, we will computationally estimate the true (unobservable) value distribution across players of the object(s) being auctioned (e.g., liquor licenses). The estimation method used will be non-parametric "distance minimization" between the observed out-of-sample distribution of bids, and the predicted out-of-sample distribution of bids using optimally calibrated parameter values. E.g., Athey, Levin, and Seira (QJE 2011) use their estimated model to make comparative static predictions and test that for fit against data from timber auctions.

Finally, to analyse which auction design would have best met the various aims of the auction designer, we will use the calibrated model, parameters, and the estimated valuations of the bidders. In particular, using these we will simulate the revenue, efficiency, and other metrics of importance for different auction designs. In addition to the use of simulation described above, to analyse alternate auction designs, we will use simulations of variations of the estimated model, parameters (like risk aversion, budgets etc.), and value distributions to analyse the different rates with which different auction designs can meet the various possible aims of the auction designer.

RESEARCH PROJECTS

Ongoing Projects:

Can an AI Coach Help You Lose More Weight Than a Human Coach: Empirical Evidence From a Mobile Fitness Tracking App
Prof. Anuj Kapoor

High-frequency trading: Measuring latency from big data
Prof. Anirban Banerjee

Causes, Symptoms and Consequences of Sociocultural polarization
Prof. Samrat Gupta

Employee Reviews - A Text Mining Perspective
Prof. Adrija Majumdar

When A Machine Knows When You Are Happy (vs. Upset)
Prof. Hyokjin Kwak

Models of implied volatility and information content of option prices
Prof. Sobhesh Kumar Agarwalla and Prof. Vineet Virmani

Hiring for the Future – A People Analytics Approach
Prof. Aditya Christopher Moses

An iterative gradient-based bilevel approach for hyperparameter tuning in machine learning
Prof Ankur Sinha

CONDUCTED SEMINARS

CDSA Webinar
Youtube channel

Advertising as Information for Ranking E-Commerce Search Listings (Dec 1, 2022)
Prof. Joonhyuk Yang (University of Notre Dame)

Smiles in Profiles: Improving Fairness and Efficiency Using Estimates of Users Preferences in Online Marketplaces (Nov 21, 2022)
Dr. Emil Palikot (Stanford Graduate School of Business)

Mechanisms to Infer the Wisdom of the Crowd (Oct 31, 2022)
Prof. Mallesh Pai (Rice University)

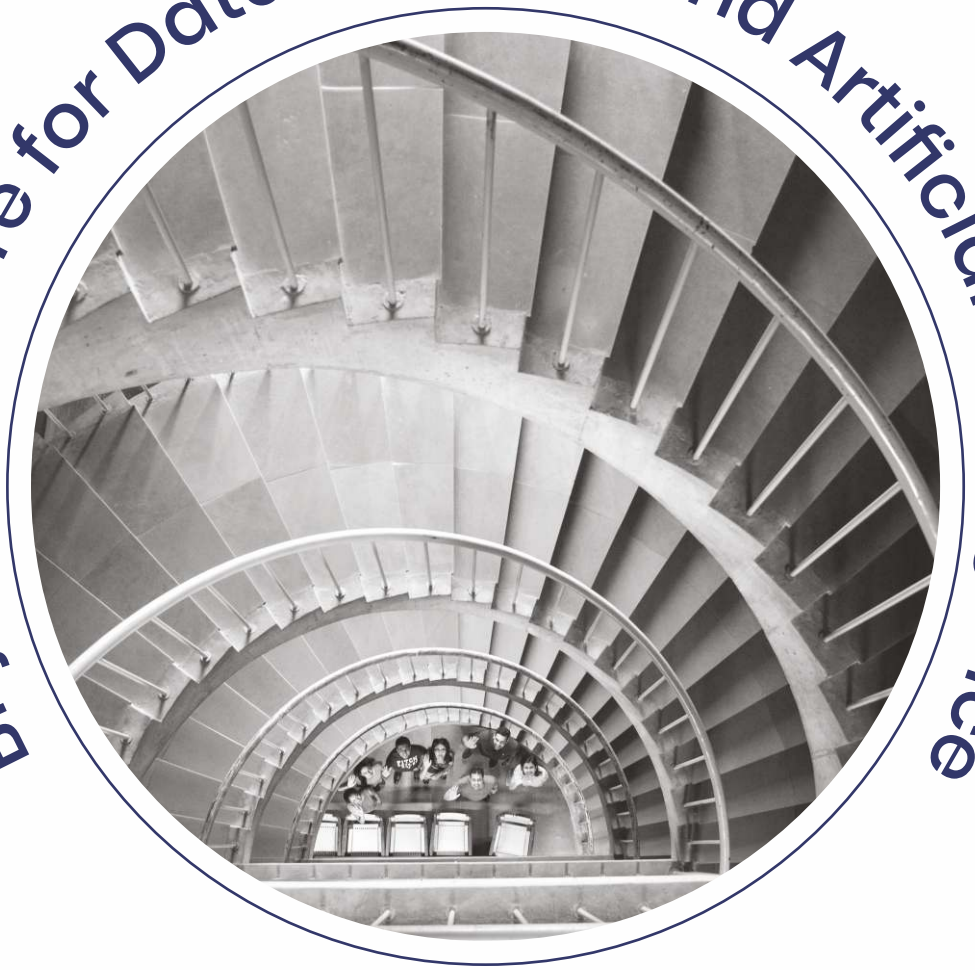
Search Neutrality and Competition between First-Party and Third-Party Sellers (Sept 19, 2022)
Prof. Bobby Zhou (University of Maryland)

Scalable Bundling via Dense Product Embeddings (Sept 12, 2022)
Dr. Madhav Kumar (MIT Sloan School of Management)

You cannot optimize what you do not know: Evidence from a Reform (Aug 25, 2022)
Prof. Avner Strulov-Shlain (University of Chicago)

Misinformation in the context of COVID-19 Pandemic: An investigation of health harm characteristics and related social media conversations (Aug 23, 2022)
Prof. H. R Rao (University of Texas)

Brij Disa Centre for Data science and Artificial Intelligence



IIMA

विद्याविनियोगाद्विकासः

भारतीय प्रबंध संस्थान अहमदाबाद

INDIAN INSTITUTE *of* MANAGEMENT AHMEDABAD

वस्त्रापुर, अहमदाबाद ३८००१५, भारत । Vastrapur, Ahmedabad 380015, India.

+91 79 7152 7742 cdsa@iima.ac.in www.iima.ac.in/cdsa/

Follow us on Linkedin

Co-Chairpersons:

Prof. Ankur Sinha (email: asinha@iima.ac.in) & Prof. Anindya S. Chakrabarti (email: anindyac@iima.ac.in)

Centre Head: Debjit Ghatak (email: gm-cdsa@iima.ac.in) | Centre Secretary: Anjali Nair (email: cdsa@iima.ac.in)