Math. Appl. 15(2s)2026, 54–75

Text Mining and Social Media Sentiment Analysis Measuring Brand Equity and Customer Satisfaction Post-Product Launch

Mahesh Reddy Challa[1]

Abstract: *The weeks after a product launch are when a brand's promise meets customer experience, and social media is the only source that records that encounter as it happens. Firms increasingly turn posts and reviews into sentiment indicators, yet the link between a classifier's accuracy on individual posts and the quality of the brand and satisfaction indicators built from them is rarely examined. This paper studies that link with a text-mining pipeline that covers spam filtering, sentiment classification, aspect-level aggregation, advocacy mining and topic discovery, evaluated on a simulated corpus of 18,560 launch-period posts over 12 weeks in which the true customer experience, including a delivery failure, a firmware fix and a price promotion, is known. A rule-based lexicon reaches only 0.534 macro-F1 on posts from weeks 4 to 12, whereas supervised linear models trained on 3,000 labelled posts from the first three weeks reach 0.795 (naive Bayes), 0.811 (logistic regression) and 0.815 (linear support vector machine). Despite that gap, the weekly net sentiment series of every method follows the true trajectory of customer satisfaction closely (correlations of 0.89 to 0.95), but the methods differ in accuracy of level: mean absolute error against the true net sentiment is 0.021 for the support vector machine and 0.050 for the lexicon, which compresses the delivery dip and the later recovery. Aspect-level tracking recovers the delivery, battery and price episodes (correlations of 0.94, 0.88 and 0.86) but is weaker for support (0.65). A brand-equity composite built from net sentiment, advocacy phrases and mention volume tracks the simulated brand equity at 0.96, whereas mention volume alone does not track it at all. Periodic relabelling brought no measurable gain in this corpus, and unsupervised topics aligned only modestly with product aspects. All data are synthetic, so the results illustrate an evaluation design and should not be read as expected performance on real platforms.*

Introduction

A product launch compresses a great deal of uncertainty into a few weeks. Marketing has built expectations, supply chains are tested at volume, and customers are for the first time using something they have only seen advertised. What they say about the experience, in posts, reviews and replies, is visible to millions of other potential buyers and shapes the brand's reputation long before any quarterly survey is fielded. Social media has therefore become a standard source for monitoring brand health and customer satisfaction after a launch, and firms routinely convert streams of text into sentiment dashboards [1], [2].

[1]3151 Solana circle Denton Texas zip code 76203.

Keywords: text mining, sentiment analysis, social media, brand equity, customer satisfaction, product launch, aspect-based sentiment, topic modelling, social listening.

The idea has strong intellectual foundations. Brand equity, understood as the differential effect of brand knowledge on customer response, depends on what customers know and feel about the brand [3], [4]. Customer satisfaction arises from the comparison between expectations and perceived performance [5]. Electronic word of mouth carries both elements at scale and influences the behaviour of others [6], [7]. Sentiment analysis, the computational study of opinions expressed in text, offers a way to turn that word of mouth into numbers [8], [9].

In practice, however, three difficulties make the numbers less trustworthy than the dashboards suggest. First, text is hard: sarcasm, negation, mixed opinions, implicit complaints with no opinion words, slang and misspellings all reduce classifier accuracy, and accuracy on a benchmark dataset does not transfer reliably to a new product and platform [10]. Second, social media samples are not representative: people who post are not typical customers, negative experiences may be over-represented, and automated accounts and promotional posts add noise [11], [12]. Third, and least discussed, individual-level accuracy is not the same as indicator-level validity. A classifier that is wrong on a fifth of posts may still produce a faithful weekly trend if its errors are stable, whereas a classifier with moderate accuracy may distort the level of the indicator in a way that matters for targets and thresholds. What a manager needs to know is whether the weekly series moves with the true customer experience, by how much, and which aspects of the product are driving it.

This paper addresses these questions with an evaluation that compares text-mining methods on two levels, the level of the individual post and the level of the weekly indicator, in a setting where the truth at both levels is known. We build a pipeline that covers spam filtering, sentiment classification, aspect-level aggregation, advocacy phrase mining and topic discovery, and we evaluate it on a simulated launch in which customer experience changes through three events: a delivery failure in the third and fourth weeks, a firmware update that improves battery life from the sixth week, and a price promotion from the eighth week.

The contributions are as follows.

1. A text-mining pipeline for post-launch monitoring that produces sentiment, aspect-level and brand-equity indicators from raw posts.
2. A two-level evaluation protocol that distinguishes post-level classification quality from indicator-level validity, with bootstrap uncertainty for the former and correlation and error measures against known truth for the latter.
3. Evidence from a controlled corpus on how four classifiers compare, how much the weekly indicators depend on classifier accuracy, which brand-equity proxies work, and how labelling effort and relabelling affect performance.
4. Practical guidance on what to validate before trusting a social-media sentiment dashboard.

The data in this study are synthetic. They were generated to include the difficulties listed above, but they cannot reproduce the full variety of real language, and the numbers should be treated as an illustration of the evaluation design.

LITERATURE REVIEW

A. Brand equity and customer satisfaction

Brand equity has been conceptualised in several ways, but a common thread is that it resides in the minds of customers. Aaker described it as a set of assets linked to a brand, including awareness, perceived quality and loyalty [3]. Keller defined customer-based brand equity as the differential effect of brand knowledge on consumer response to marketing, and stressed that it depends on brand awareness and on the favourability, strength and uniqueness of brand associations [4]. Satisfaction is treated as a related but distinct construct. In the expectation-disconfirmation model, satisfaction results from comparing perceived performance with prior expectations, and it influences later attitudes and repurchase [5]. A single loyalty-oriented question, the likelihood of recommending a company, has been promoted as a simple indicator of growth [13]. These constructs matter after a launch because early experience feeds both satisfaction and the brand associations that will carry forward.

B. Electronic word of mouth

Consumers use online opinions to inform their decisions, and the volume and valence of those opinions influence sales. Studies of opinion platforms show that consumers are motivated to post by concern for others, desire for social interaction and self-enhancement among other reasons [6]. Analysis of online book reviews found that the effect of reviews on sales depends on their content and ratings, with negative reviews often having more impact than positive ones [7]. The implication for monitoring is that posts are both a measure of customer experience and a cause of future outcomes, so a launch that produces visible negative word of mouth may suffer beyond the immediate dissatisfaction.

C. Sentiment analysis and opinion mining

Sentiment analysis has been studied extensively since the early 2000s. Surveys describe tasks at the level of documents, sentences and aspects, as well as approaches based on lexicons, classical machine learning and, more recently, neural networks [8], [9]. An early landmark study applied naive Bayes, maximum entropy and support vector machines to movie reviews and found that machine learning methods outperformed simple keyword counts [14]. Lexicon-and-rules tools designed for social media use dictionaries of sentiment words with intensity scores and rules for negation and emphasis, and perform well on short informal text for what they were designed [15]. Pre-trained transformer language models have since raised accuracy on many benchmarks [16], but they require more resources and, like other methods, can perform worse on text that differs from their training data. Aspect-based sentiment analysis goes further by assigning sentiment to particular features of a product, which is what product teams need to act [17].

D. Challenges in social-media text

Social-media text is short, informal and context-dependent. Sarcasm, in which the literal meaning is the opposite of the intended one, is a recognised difficulty, and research on detecting it in microblogs shows that

simple lexical features are insufficient [18]. Negation, mixed opinions and implicit sentiment, such as describing an event that is obviously bad without using any opinion word, cause further errors. Misspellings, abbreviations, emoji and hashtags change vocabulary quickly. Language also drifts: new product features, events and slang appear after a model is trained, so models need monitoring and sometimes retraining.

E. Validity of social media as a measurement source

Social media data are not a random sample of customers. Researchers have warned about sampling bias, platform-specific behaviour, the presence of automated accounts and the temptation to over-interpret large volumes of data [11], [12]. Work comparing text sentiment with polls found strong correlations in some settings, with the relationship depending on how the text was filtered and aggregated [19]. For brand and satisfaction monitoring, this means that the indicators must be validated against external measures, such as surveys, and not assumed to represent the customer base.

F. Topic discovery

Unsupervised topic models and matrix factorisation methods discover recurring themes in text without labels [20]. They are used in practice to find what customers talk about, including themes that were not anticipated. Their output requires interpretation, and topics often reflect style and phrasing as well as subject matter.

RELATED WORK

Sentiment classifiers on short text. Comparative studies typically report results on public datasets, mostly of reviews or microblogs, and rank lexicon methods, classical classifiers and neural models [8], [9], [14], [15], [16]. The consistent finding is that supervised methods trained on in-domain labels outperform generic lexicons, and that the margin shrinks when the training data are small or differ from the target domain. Our study follows this tradition for the post-level comparison but extends it to the weekly indicator level.

Aspect-based analysis. Shared tasks on aspect-based sentiment analysis defined the problem of identifying aspect terms and their polarity in reviews of products and services [17]. We use a simpler design in which each post has a dominant aspect, since our emphasis is on how errors propagate to aspect-level time series and not on aspect extraction.

Social media indicators and external outcomes. Several studies relate aggregated text sentiment to external outcomes, such as political polls and consumer confidence [19], and to financial indicators. Findings are mixed and sensitive to processing choices. We do not claim any such external prediction, and instead check how well text-based indicators follow a known latent truth in a controlled setting.

Brand monitoring with social data. Practitioner and academic work uses mention volume, sentiment and engagement as proxies for brand health. The validity of each proxy is rarely tested against a ground truth. We provide such a test for volume, advocacy phrases, net sentiment and their composite.

Gap. There are few evaluations that connect post-level classification quality to the validity of the weekly indicators built from it in a setting with known

ground truth, and that examine, at the same time, the effects of spam, labelling effort, retraining and aspect granularity. This paper provides a controlled study of those questions, with the limitations of synthetic data explicitly stated.

METHODOLOGY

A. Overview of the pipeline

The pipeline in Fig. 1 has five stages. Raw posts are cleaned and screened for promotional spam. Cleaned posts are classified as positive, negative or neutral. Posts are assigned to product aspects, and sentiment is aggregated by week and by aspect. Advocacy phrases and mention volume are extracted. Finally, the components are combined into indicators of customer satisfaction and brand equity. A parallel branch applies unsupervised topic discovery to explore what customers discuss.

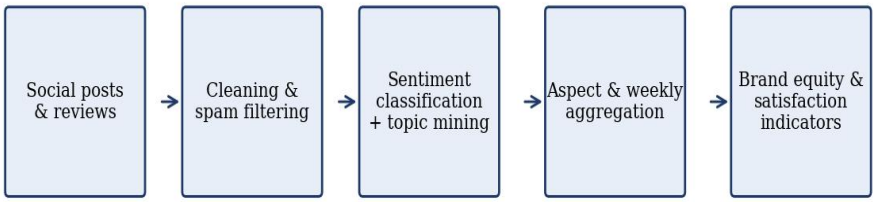


Fig. 1. Text-mining pipeline for post-launch brand and satisfaction monitoring.

B. Preprocessing and spam filtering

Text is lower-cased, and URLs, user mentions and punctuation are normalised for the lexicon method, while the supervised models also use character sequences that make them tolerant to misspellings. Promotional or automated posts, which typically contain links and discount codes and carry positive wording, would inflate sentiment if left in. We use a transparent screening rule that removes posts containing a web link or a discount-code phrase, and we compare indicators with and without it.

C. Sentiment classifiers

Four classifiers are compared.

Lexicon with rules. A dictionary of positive and negative words, with the sign reversed when a word follows a negation, strengthened after an intensifier, and with emoji counted as evidence. A post is labelled positive or negative when the total score is far enough from zero and neutral otherwise. No training data are needed.

Naive Bayes. A probabilistic classifier on counts of single words and word pairs, which is fast and a common baseline.

Logistic regression. A linear classifier on weighted word and word-pair features together with character sequences of two to five characters, which provide robustness to spelling variation.

Linear support vector machine. A linear maximum-margin classifier on the same features.

Supervised models are trained on 3,000 labelled posts sampled from the first three weeks after launch, which mimics a team that labels an initial sample and then deploys the model. Six per cent of the training labels are replaced by random labels to mimic annotator error. The models are then applied to posts from the fourth week onward, so that evaluation includes language that appears after training, such as event-related vocabulary.

D. Aspect assignment and aggregation

Each post is assigned to a product aspect, one of battery, price, delivery, support and design. In the simulation, the aspect is known, and the study concentrates on how sentiment errors affect aspect-level series. Sentiment is aggregated into *net sentiment*, defined as the share of positive posts minus the share of negative posts in a week, a measure commonly used in social listening because it is simple and bounded between minus one and one.

E. Advocacy mining and brand-equity composite

Advocacy is measured by the share of posts that contain recommendation phrases such as “highly recommend”, “must buy”, “worth every penny” and “would buy again”. Mention volume is the weekly count of cleaned posts and is used as a proxy for awareness. A brand-equity composite averages the standardised weekly values of net sentiment, advocacy share and mention volume, so that each component has equal weight. This composite is meant as a transparent starting point and not as an optimal index.

F. Topic discovery

Non-negative matrix factorisation is applied to weighted word frequencies of cleaned posts, after removal of common words and the brand name, to discover five topics. Each post is assigned to the topic with the highest weight, and topic quality is assessed by comparing topics with the true aspects through purity, the share of posts that belong to the majority aspect of their topic. A purity of 0.2 would be expected by chance with five equal aspects, and a purity close to one would indicate topics that coincide with aspects.

G. Evaluation design

Post-level. Classifiers are compared on posts from weeks 4 to 12 using accuracy, per-class F-scores and macro-F1, the unweighted mean of the F-scores of the three classes, which does not favour the majority class. Uncertainty is quantified by bootstrap resampling of test posts (200 resamples). We also report accuracy on negated and sarcastic posts.

Indicator-level. For each classifier, the weekly net sentiment is compared with two truths: the net sentiment computed from the true post labels, and the true latent customer satisfaction used to generate the corpus. We report the Pearson correlation across the 12 weeks, which measures whether the series moves together with the truth, and the mean absolute error against the true net sentiment, which measures level accuracy. Because only 12 weekly points are available, correlations are descriptive and carry considerable uncertainty.

Learning-effort. We vary the number of labelled training posts and test

whether adding 1,000 labelled posts from weeks 6 and 7 improves performance on weeks 8 to 12.

EXPERIMENTAL SETUP

A. Why synthetic data

Real launch data with ground truth are not available: the true level of customer satisfaction cannot be observed from posts, labels are expensive and subjective, and the content of real platforms cannot be shared without privacy and terms-of-service concerns. A synthetic corpus lets us know exactly what customers experienced and what each post expresses, so that errors can be attributed. Its limitation is realism. The language generator, however varied, is far simpler than human writing, and some conclusions, in particular those about sarcasm, are affected.

B. Corpus generation

The simulated launch involves a device with five aspects. Customer satisfaction with each aspect follows a trajectory. Battery satisfaction starts around 0.52 and rises to about 0.72 following a firmware fix at the sixth week. Delivery satisfaction starts at 0.70, falls to about 0.39 in the third and fourth weeks during a failure, and recovers. Price satisfaction rises from 0.50 to 0.62 with a promotion from the eighth week. Support satisfaction declines slowly from 0.60 to 0.48, and design stays at 0.80. The overall satisfaction index is a weighted average of the aspects.

Each week, between about 1,300 and 2,400 posts are produced, with a launch peak in the first two weeks and a smaller peak at the promotion. Each post concerns one aspect and has a true sentiment label drawn from the satisfaction of that aspect. Posts are generated in four styles: explicit opinion (45%), implicit sentiment with no opinion word, such as “stopped working after five days” (27%), mixed opinions with both positive and negative words (14%), and expectation statements with contrasting words, such as “expected something great but got terrible” (14%). Implicit posts are drawn from templates of the target class with 72% probability and from another class otherwise, so they are inherently ambiguous. About one in ten negative posts is sarcastic, using positive words, and some positive posts begin with “oh great” while being sincere. Twenty per cent of posts are long reviews with an additional neutral remark, half of posts carry a hashtag, 7% of long words are misspelled, and new vocabulary linked to events appears after they occur. Negative posts are 30% more likely to be written than a satisfaction level alone would imply, simulating the tendency of dissatisfied customers to speak up. Five per cent of additional posts are promotional spam with links and discount codes.

The corpus contains 18,560 posts, of which 47.6% are positive, 34.2% negative and 18.3% neutral among non-spam posts, and 4.8% of all posts are spam. The training sample contains 3,000 non-spam posts from weeks 1 to 3. The test set is the non-spam posts from weeks 4 to 12, roughly 12,200 posts.

C. Software

All methods were implemented in Python with the scikit-learn library [21],

and all random seeds are fixed. The full set of experiments runs in about one minute on a single processor.

D. Experiments

We report seven experiments: post-level classification (E1), the effect of adding labelled data from later weeks (E2), indicator-level tracking with and without spam filtering (E3), aspect-level tracking (E4), brand-equity proxies (E5), topic discovery (E6), and the effect of training-set size (E7)..

RESULTS

A. Post-level classification

Table 1 reports performance on posts from weeks 4 to 12, and Fig. 2 plots macro-F1 with bootstrap intervals.

Table 1. Post-level classification on weeks 4 to 12.

Method	Accuracy	Macro-F1 (95% interval)	F1 negative	F1 neutral	F1 positive	Accuracy on negated posts	Accuracy on sarcastic posts
Lexicon with rules	0.531	0.534 (0.526 to 0.542)	0.545	0.447	0.611	0.941	0.025
Naive Bayes	0.804	0.795 (0.786 to 0.803)	0.773	0.774	0.836	0.995	1.000
Logistic regression	0.828	0.811 (0.803 to 0.820)	0.809	0.758	0.866	1.000	1.000
Linear SVM	0.831	0.815 (0.807 to 0.824)	0.814	0.762	0.868	1.000	1.000

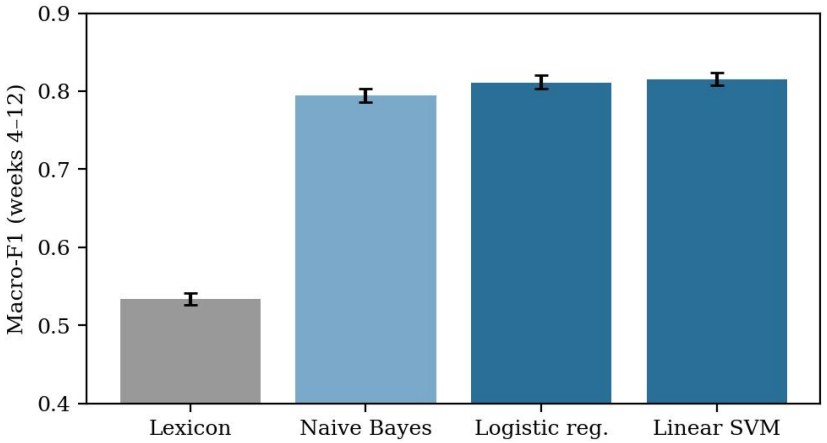


Fig. 2. Macro-F1 of the four classifiers on weeks 4 to 12 (error bars: 95% bootstrap intervals).

The supervised linear models clearly outperform the lexicon, with macro-F1 values between 0.795 and 0.815 against 0.534. The bootstrap intervals of logistic regression and the support vector machine overlap almost completely, while the interval of naive Bayes only just touches that of logistic regression, so naive Bayes is slightly lower. The lexicon fails mainly

on implicit sentiment and mixed posts, since it can only count opinion words, and the supervised models learn that phrases such as “stopped working” or “bought a second one” carry sentiment even though they contain no opinion word. The neutral class is the hardest for the supervised models (F1 between 0.758 and 0.774), a typical pattern, since neutral posts are close to the other classes in the ambiguous styles.

Negation is handled well by all methods in this corpus, which reflects the simple negation constructions in the generator. Sarcasm shows the starkest contrast: the lexicon labels only 2.5% of sarcastic posts correctly, because the sentiment words are positive and the true label is negative, whereas the supervised models label all of them correctly. This perfect score should be treated with caution. The sarcastic posts in the generator come from four templates with distinctive phrases, which the supervised models memorise, and real sarcasm is far more varied and context-dependent. The result demonstrates that supervised models can learn recurring sarcastic idioms from labelled data, but not that they solve sarcasm.

B. Indicator-level validity

Table 2 compares the weekly net sentiment series produced by each method, after spam filtering, with the truth.

Table 2. Weekly net sentiment versus ground truth (12 weeks).

Method	Correlation with true net sentiment	Correlation with true satisfaction	Mean absolute error, spam filtered	Mean absolute error, no filter
Lexicon with rules	0.954	0.949	0.050	0.041
Naive Bayes	0.940	0.887	0.033	0.028
Logistic regression	0.967	0.926	0.024	0.046
Linear SVM	0.968	0.922	0.021	0.045

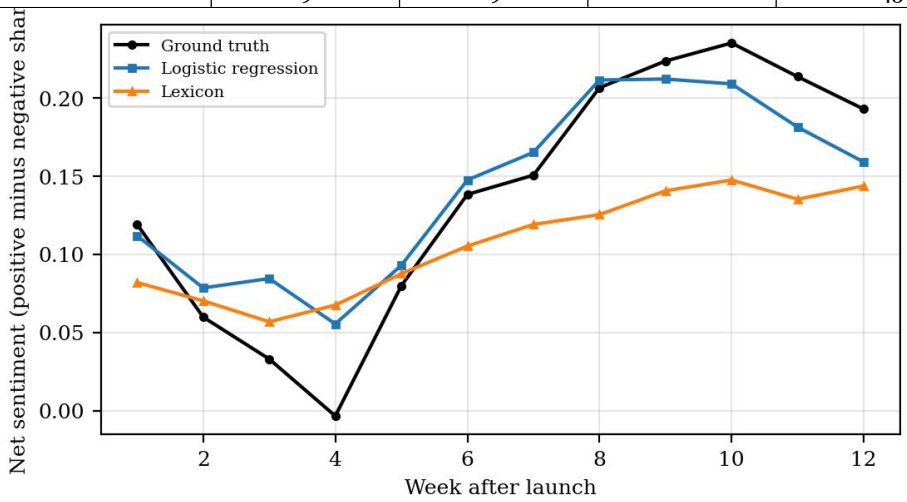


Fig. 3. Weekly net sentiment: ground truth against two methods.

The first point is that every method tracks the direction of change well. Correlations with the true net sentiment lie between 0.94 and 0.97, and correlations with true satisfaction between 0.89 and 0.95, despite the large difference in post-level accuracy. The lexicon, whose post-level macro-F1 is 0.534, has a correlation with true satisfaction (0.949) that is actually the highest of the four. The reason is that its errors are systematic and fairly

stable across weeks, so that the series is a compressed copy of the truth. This is an important qualification of the common practice of ranking tools by post-level accuracy alone: for the question of whether the series rises and falls with customer experience, a crude method may be adequate.

The second point is that level accuracy differs sharply. The mean absolute error against the true net sentiment is 0.021 for the support vector machine and 0.024 for logistic regression, but 0.050 for the lexicon. Fig. 3 shows why: the lexicon's series ranges only from about 0.06 to 0.15, whereas the true series ranges from about zero at the bottom of the delivery failure to about 0.24 at its peak. The lexicon compresses the trough and the recovery. For a manager who wants to know whether net sentiment has fallen below a threshold, or by how much it has changed, this compression matters even though the trend direction is right. The supervised series follows the truth with errors of two to three points of net sentiment, although it still overestimates the trough in week 4 (0.056 against a true value of minus 0.003).

The third point concerns spam. Removing promotional posts reduces the error of the supervised models by about half (from 0.045 or 0.046 to 0.021 or 0.024), as spam adds positive posts that were not part of the customer experience. For the lexicon and naive Bayes, filtering did not help, and slightly increased the error, because their series were biased downward, so that the extra positive posts happened to offset part of the bias. This is an accident of compensating errors and not a reason to leave spam in. The spam screening rule worked perfectly in this corpus because the synthetic spam always contains a link or a code; real spam is more evasive, so this result overstates what a simple rule can do.

A fourth observation is that the relationship between net sentiment and satisfaction is a matter of scale. True net sentiment ranges from about zero to 0.24, whereas true satisfaction ranges from 0.55 to 0.68. Net sentiment is lower than satisfaction would suggest because negative customers are more likely to post. The shape of the series is informative but its absolute level is not a direct estimate of satisfaction, and a mapping calibrated on survey data would be needed to turn net sentiment into a satisfaction score.

C. Aspect-level tracking

Table 3 reports the correlation between the weekly aspect-level net sentiment from the logistic regression and the true aspect satisfaction.

Table 3. Aspect-level tracking with logistic regression.

Aspect	Event in the simulation	Correlation with true aspect satisfaction
Delivery	Failure in weeks 3 and 4, then recovery	0.94
Battery	Firmware fix from week 6	0.88
Price	Promotion from week 8	0.86
Support	Slow decline	0.65
Design	No change	Not defined (constant)

The three aspects with clear events are recovered with correlations from 0.86 to 0.94, which means that the pipeline would have alerted a product team to the delivery failure, confirmed the effect of the firmware fix and registered the effect of the promotion. The slow decline in support satisfaction is recovered less well (0.65), since a gradual change of about one point of satisfaction per week is small compared with sampling noise at about 200 posts per aspect per week. This suggests that aspect-level

monitoring is better at detecting sharp events than slow drifts, and that detecting drifts requires larger samples or smoothing over longer windows.

D. Brand-equity proxies

Table 4 compares the proxies of brand equity with the simulated brand equity, which combines satisfaction, advocacy and awareness.

Table 4. Correlation of social-media proxies with simulated brand equity (12 weeks).

Proxy	Correlation
Mention volume alone	−0.00
Advocacy phrase share alone	0.84
Composite of net sentiment, advocacy and volume	0.96
Composite against true satisfaction	0.80

Mention volume, a common proxy for awareness and buzz, has no relationship with the simulated brand equity. The volume is dominated by the launch peak in the first two weeks, when buzz is high but brand equity is still forming, and by the promotion peak, which raises volume without a corresponding change in the underlying experience. Advocacy phrases correlate at 0.84 and the composite at 0.96. The composite is naturally close to the simulated brand equity because both include advocacy and volume-related components by construction, so this high value is partly built in and should not be read as evidence of external validity. A more informative result is the negative one: volume alone is a poor proxy. The composite correlates only 0.80 with satisfaction, and the net sentiment alone correlates at 0.93, which suggests that a composite helps with brand equity but dilutes the signal for satisfaction.

E. Topic discovery

Five topics discovered by matrix factorisation were assigned to posts and compared with the true aspects. The purity was 0.35, against a chance level of about 0.2 to 0.25. The topics reflected style and phrasing as well as aspects. One topic grouped delivery, design, price and support words around the word “expected”, corresponding to the contrast style, and another grouped implicit phrases such as “bought a second one for my sister”, across aspects. Battery appeared in two topics and delivery in three. For an analyst, the topics would reveal that customers talk about both specific features and ways of expressing experience, but they would not by themselves provide a clean aspect breakdown. Unsupervised topics are therefore best used for exploration, with aspect classification based on dictionaries or supervised models for tracking.

F. Labelling effort and relabelling

Fig. 4 shows macro-F1 of logistic regression as a function of the number of labelled posts.

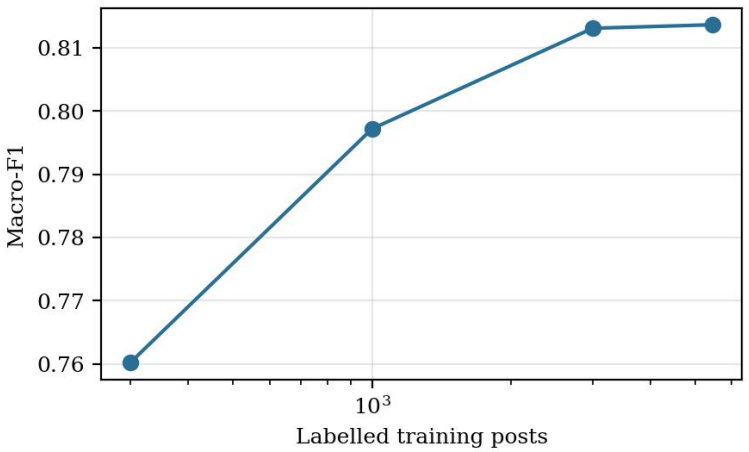


Fig. 4. Macro-F1 of logistic regression by number of labelled training posts.

Performance rises from 0.760 with 300 labelled posts to 0.797 with 1,000 and 0.813 with 3,000, and is flat beyond that (0.814 with 5,459). Most of the benefit therefore comes from the first thousand labels, and the returns diminish quickly. The ceiling of about 0.81 is set by the ambiguity built into the corpus (mixed posts, ambiguous implicit posts and annotation noise) and not by lack of data.

Adding 1,000 labelled posts from weeks 6 and 7 changed macro-F1 on weeks 8 to 12 only marginally: from 0.792 to 0.799 for naive Bayes, from 0.810 to 0.811 for logistic regression and from 0.8137 to 0.8139 for the support vector machine. The corpus contains vocabulary drift, but it is mild, since new event-related terms appear in a small share of posts and the models rely on many other cues. In real settings, drift may be stronger, for instance when a new competitor, a recall or a viral meme changes the language, and the conclusion that relabelling is unnecessary should not be generalised. A sensible practice is to monitor performance on a small, freshly labelled sample each week and retrain when it deteriorates.

DISCUSSION

A. Interpretation

The study supports four messages. First, supervised classifiers trained on a few thousand in-domain labels are far better than a generic lexicon at the level of individual posts, mainly because they can learn implicit and idiomatic expressions of sentiment. Second, post-level accuracy is a poor guide to the usefulness of the weekly indicator for trend detection, since all methods followed the direction of customer experience closely. Third, post-level accuracy matters for the level and magnitude of the indicator: the lexicon compressed the delivery dip and recovery, which would have understated the severity of the delivery problem. Fourth, the proxies matter: mention volume was uninformative about brand equity, advocacy phrases and net sentiment were informative, and the composite, while attractive, mixes signals with different dynamics.

B. Implications for practice

Validate indicators, not only classifiers. Teams should evaluate a sentiment tool by asking whether its weekly series reproduces known events and external measures such as survey scores, and not only by the accuracy on a benchmark. A small set of weekly labelled posts and a periodic comparison with a customer survey provide the necessary checks.

Use a supervised model once a few hundred labels exist. The learning curve shows that 1,000 labelled posts bring most of the benefit. A generic lexicon can serve as a starting point but should be replaced when labels are available, particularly if the product domain has its own vocabulary.

Report levels carefully. Net sentiment is a relative indicator. Because negative customers tend to post more, its level will be lower than satisfaction suggests. Targets should be set on changes relative to a baseline or after calibration against surveys, not on the raw level.

Filter spam, but do not rely on a simple rule. Promotional and automated posts distort indicators, and the effect on the supervised series was substantial. Production systems should use classifiers or platform signals for bot and spam detection, and monitor the share of removed posts.

Choose proxies with care. Mention volume reflects marketing activity and launch buzz and should not be used as a stand-in for brand equity. Advocacy and sentiment are more informative. Composites should be validated against brand tracking surveys before being used for decisions.

Look at aspects. The aspect-level series told a product team where the problems and improvements were, which a single overall score cannot do. Slow drifts need larger samples and longer windows.

C. Ethical and privacy considerations

Social posts are personal data in many jurisdictions, even when public. Monitoring should comply with platform terms and data protection law, minimise the retention of identifying information, and avoid linking posts to individuals without a legitimate basis. Sentiment models can have unequal accuracy across dialects, languages and communities, which can lead to some groups being systematically misread. Teams should evaluate performance across such groups where data permit, and avoid using sentiment indicators for decisions about individuals.

D. Threats to validity

Synthetic data. The principal limitation is that the corpus is synthetic. Real language is more varied, sarcasm and context are harder, labels are more subjective, and customer experience is not known. The numbers therefore illustrate the evaluation design, and the ranking of methods on real data may differ, in particular for neural language models, which were not evaluated.

Templated sarcasm and negation. The perfect accuracy of supervised models on sarcastic and negated posts reflects the limited number of templates. We do not claim that these phenomena are solved.

Simple spam. The spam in the corpus is easy to identify. Real spam and coordinated campaigns are harder to detect.

Defined brand equity. The simulated brand equity is a weighted combination of satisfaction, advocacy and awareness defined by us. The high correlation of the composite with it is partly circular. An external

validation against real brand tracking data is needed.
Few time points. Weekly correlations rest on 12 points and have wide uncertainty. They are descriptive.
Single corpus and a small set of models. We evaluated one corpus and four classifiers. Transformer models, ensembles and aspect extraction models might change the conclusions about post-level accuracy, though less so, we expect, the conclusion that indicator-level validity depends on stability of errors.

E. Future work

Immediate priorities are validation on real launch data with survey benchmarks, evaluation of pre-trained transformer models alongside the classical methods, and calibration of net sentiment to satisfaction scores. Further directions include automatic aspect extraction, detection of coordinated inauthentic activity, multilingual and code-mixed text, handling of images and video, and methods for early warning, such as change detection on aspect-level series with controlled false alarm rates.
.

CONCLUSION

This paper examined how text mining and sentiment analysis can be used to measure brand equity and customer satisfaction after a product launch, and in particular how the quality of classification of individual posts relates to the validity of the indicators built from them. In a synthetic corpus of 18,560 posts over 12 weeks with known customer experience, supervised linear models reached macro-F1 between 0.795 and 0.815 against 0.534 for a lexicon. All methods tracked the direction of satisfaction closely, but the lexicon compressed the size of the delivery dip, with an error in net sentiment of 0.050 against 0.021 for the best supervised model. Aspect-level series recovered the delivery, battery and price events, mention volume was useless as a brand-equity proxy, a composite including advocacy and sentiment was informative, and most of the benefit of labelling came from the first thousand labelled posts.
The practical message is to validate the indicator and not only the classifier: check that the series reproduces known events, calibrate its level against surveys, filter promotional content, examine aspects and choose proxies with care. Because the data are synthetic, the results should be taken as a demonstration of an evaluation design. The next step is to apply the same two-level protocol to real launch data and survey benchmarks.

References

[1] Pawan Kumar Singh, Scaling Gate-All-Around (GAA) Transistor Architectures for Sub-2nm AI Accelerators Using Atomically Thin 2D Materials, Vol. 33 No. 08 (2024): JOCAAA, Journal Name: Journal of Computational Analysis and Applications, Journal's ISSN: 1521-1398 (Paper),1572-9206 (Online), https://eudoxuspress.com/index.php/pub/article/view/5713
[2] Pawan Kumar Singh, Energy-Efficient Task Scheduling in Green Cloud Computing Using Neuromorphic Spiking Neural Networks, Vol. 33 No. 08 (2024): JOCAAA, Journal Name: Journal of Computational Analysis and

Applications, Journal's ISSN: 1521-1398 (Paper),1572-9206 (Online), https://eudoxuspress.com/index.php/pub/article/view/5712
[3] Rangavinay Teja Royal Jagga, 2026, DOI: 10.1109/GCWCN66157.2025.11448495, Date of Conference: 22-23 November 2025, Date Added to IEEE Xplore: 25 March 2026 ,https://ieeexplore.ieee.org/document/11448495
[4] Rangavinay Teja Royal Jagga , 2026, Subtle Data Contamination in Visual AI: Covert Model Backdoors via Resampling Exploits, Date of Conference: 22-23 November 2025, Date Added to IEEE Xplore: 25 March 2026 ,https://ieeexplore.ieee.org/document/11448514
[5] Rangavinay Teja Royal Jagga, 2026, Robotic Tactile-Vision Integration for Interactive Object Disentanglement, https://ieeexplore.ieee.org/document/11448337
[6] Rangavinay Teja Royal Jagga, 2026, A Discounted Future Reward Framework for Intelligent Loyalty Optimization using CLTV-Aware Segmentation and Churn Prediction, https://ieeexplore.ieee.org/document/11497560
[7] Anjaneyulu Gudla; Manish Kumar; Vamshi Jeksani; Kiran Kumar Reddy; Gayathri Band, Harnessing Artificial Intelligence Cybersecurity: Cutting Edge Collaboration of SAP AWS to Safeguard Life Science Data. 2026 https://doi.org/10.1109/aiei69164.2026.11497833
[8] G. Mahender, Department of CSE(Data Science), Vardhaman College of Engineering ; Vamshi Jeksani; Manish Kumar; Koppuravuri Sri Lakshmi Sruthi; Kiran Kumar Reddy, SAP Cloud System on AWS for Risk Detection and Integration Artificial Intelligence Scalable Cybersecurity 2026 https://doi.org/10.1109/aiei69164.2026.11497435
[9] Kiran Kumar Reddy; Kamalakar Gunupati; Manish Kumar; Pell Reddy Rajender Reddy; Ramesh Julakanti; Ram Reddy Jonnalagadda, SAP System Optimization Using AI-Driven Process Automation and Predictive Modeling Maintenance for Enhanced Business Efficiency. 2025 https://doi.org/10.1109/computingcon64838.2025.11376613
[10] Amilineni, M.V. 2025. Audit-Safe Agentic RAG Framework for Regulated Enterprise Systems: A Trustworthy AI Architecture for SAP, Finance, Healthcare, and Government Workflows. INTERNATIONAL JOURNAL OF ADVANCES IN SIGNAL AND IMAGE SCIENCES. (Jan. 2025), 34–46. DOI: https://doi.org/10.29284/7fjzk883
[11] Pawan Kumar Singh, ACCELERATING LARGE LANGUAGE MODEL TRAINING USING HYBRID QUANTUM-CLASSICAL VARIATIONAL AUTOENCODERS WITH INTEGRATED PHOTONIC QRAM, Vol. 54 No. 3 (2026): July-September 2026, Power System Protection and Control, ISSN-1674-3415, https://pspac.info/index.php/dlbh/article/view/501, DOI: https://doi.org/10.46121/pspc.54.3.45
[12] Pawan Kumar Singh , RESILIENCE ASSESSMENT OF MODERN POWER GRID INFRASTRUCTURE AGAINST EXTREME WEATHER USING PHYSICS-INFORMED NEURAL NETWORKS, Vol. 54 No. 1 (2026): January-March 2026, Power System Protection and Control, ISSN-1674-3415,https://pspac.info/index.php/dlbh/article/view/502, DOI: https://doi.org/10.46121/pspc.54.1.64
[13] Pawan Kumar Singh, MITIGATING THE EXABYTE STORAGE CRISIS: ENHANCING READ/WRITE EFFICIENCY IN 5D OPTICAL GLASS STORAGE USING SILVER NANOPARTICLE DOPING, Vol. 38 No. 11s (2025), International Journal of Applied Mathematics (IJAM) , pISSN 1311-1728, eISSN 1314-8060 ,https://ijamjournal.org/ijam/publication/index.php/ijam/article/view/2153
[14] Mr. Pawan Kumar Singh (Author), Miss. Pari Nidhi Singh

(Author) ,Intelligent Bot, ASIN : B0DBGCQY8X,Publisher : Namya Press, Publication date : November 25, 2022,Language : English, Print length : 105 pages,ISBN-10 : 9390124867,ISBN-13 : 978-9390124862,Item Weight : 7.5 ounces, Dimensions : 6.24 x 0.43 x 9.24 inches,https://www.amazon.com/Intelligent-Bot-Pawan-Kumar-Singh/dp/B0DBGCQY8X/ref=sr_1_5?crid=2VAXN5BRN5SMS&dib=eyJ2IjoiMSJ9.v3OCQvHBpy20RIjfMWhoqbOl-RLOlUJIC8Xa0Pjeqilvo2sbZ_suN3LbSCw_OA_ZD_4lgjRLMGlBwZQQ1ehZEfDpt_MJiOUfVuqqzMav0I5_mb6pN5DQ0g0U7SA826n1DQQ3QAtWwopHy4lC5IfhZXttKj9faiyATwhdtDATSSV1-r7FmzM2VQyIA0jxcfq6E6hp9ECfH1fbyoli9kUfGpLxMq4GktlzUIJO5ggrhsQ.G1AjQiVqU4HFsnSczvz-X6NXCNs3EFSn-RI8mcWWGOc&dib_tag=se&keywords=Pawan+Kumar+Singh%2C+book&qid=1787815741&sprefix=pawan+kumar+singh%2C+bo%2Caps%2C450&sr=8-5

[15] Mr. Pawan Kumar Singh (Author), STORAGE CRISES DUE TO RISE OF AI: A Comprehensive Guide to Managing the Storage Challenges in the Age of Artificial Intelligence, ASIN : B0HFK12ZCT
Publisher : EB1 Profile, Accessibility : Learn more, Publication date : August 17, 2026,Edition : 1st
Language : English, File size : 1.7 MB, Screen Reader : Supported, Enhanced typesetting : Enabled
X-Ray : Not Enabled, Word Wise : Not Enabled ,Print length : 261 pages ,ISBN-13 : 979-8951479112,Page Flip : Enabled, https://www.amazon.com/STORAGE-CRISES-DUE-RISE-Comprehensive-ebook/dp/B0HFK12ZCT/ref=sr_1_1?crid=9MF6C4WDVJR5&dib=eyJ2IjoiMSJ9.DQ8fesBch7ELmRnilDt-YA.NBSCHLmetvnbhkxwIpOoj3s0IVsCV8y6-ZCfCP6Ul1c&dib_tag=se&keywords=Pawan+Kumar+Singh%2C+book%2C+eb1+Profile&qid=1787816040&sprefix=pawan+kumar+singh%2C+book%2C+eb1+profile%2Caps%2C420&sr=8-1

[16] Parag Bharatkumar Parikh, FRAMEWORKS FOR MITIGATING SUPPLY CHAIN DISRUPTIONS AND MATERIAL PRICE VOLATILITY IN SUSTAINABLE BUILDING CONSTRUCTION, Vol. 53 No. 2 (2025): April-June 2025 , , Power System Protection and Control, ISSN-1674-3415, https://pspac.info/index.php/dlbh/article/view/447, DOI: https://doi.org/10.46121/pspc.53.2.34

[17] Parag Bharatkumar Parikh, IMPLEMENTING DIGITAL INSPECTION PROTOCOLS FOR ENHANCED QUALITY CONTROL MANAGEMENT IN HIGH-RISE CONSTRUCTION, Vol. 54 No. 3 (2026): July-September 2026, Power System Protection and Control, ISSN-1674-3415, https://pspac.info/index.php/dlbh/article/view/439, DOI: https://doi.org/10.46121/pspc.54.3.12

[18] Parag Bharatkumar Parikh , EVALUATING THE FEASIBILITY AND EFFICIENCY OF 3D-PRINTED CONCRETE TECHNOLOGY IN SUSTAINABLE MODULAR CONSTRUCTION, Power System Protection and Control, ISSN-1674-3415, https://pspac.info/index.php/dlbh/issue/view/30, DOI: https://doi.org/10.46121/pspc.54.1.61

[19] Parag Bharatkumar Parikh, EVALUATING THE FEASIBILITY AND EFFICIENCY OF 3D-PRINTED CONCRETE TECHNOLOGY IN SUSTAINABLE MODULAR CONSTRUCTION, Vol. 54 No. 1 (2026): January-March 2026, System Protection and Control, ISSN-1674-3415, https://pspac.info/index.php/dlbh/article/view/440, DOI: https://doi.org/10.46121/pspc.54.1.61

[20] Ronakkumar shah, REDEFINING BUSINESS PROCESS MANAGEMENT: THE SYNERGY OF AI AND INTELLIGENT AUTOMATION, Vol. 54 No. 3 (2026): July-September 2026, System Protection and Control, ISSN-1674-3415, https://pspac.info/index.php/dlbh/article/view/448, DOI:

https://doi.org/10.46121/pspc.54.3.14
[21] Ronakkumar shah, OPERATIONAL ACCOUNTABILITY IN ACTION: METRICS AND MECHANISMS FOR SYSTEMIC SUCCESS, Vol. 53 No. 4 (2025): October-December 2025, System Protection and Control, ISSN-1674-3415, https://pspac.info/index.php/dlbh/article/view/452, DOI: https://doi.org/10.46121/pspc.53.4.40
[22] Ronakkumar shah, BEYOND THE BLUEPRINT: AUTOMATED WORKFLOW ENFORCEMENT AND POLICY GOVERNANCE, Vol. 53 No. 2 (2025): April-June 2025 , System Protection and Control, ISSN-1674-3415 ,https://pspac.info/index.php/dlbh/article/view/451, DOI: https://doi.org/10.46121/pspc.53.2.35
[23] Tushita Agarwal; Swaminathan Sethuraman, Mapping Customer Behavior: Visual Analysis of Online Retail Interactions, Date Added to IEEE Xplore: 17 February 2026, DOI: https://doi.org/10.1109/IC2NC67409.2025.11376261
[24] Tushita Agarwal, Efficient Data Summarization and Comparison via Grid-Based Statistical Proxies for Large-Scale Datasets, Date Added to IEEE Xplore: 03 December 2025, DOI: https://doi.org/10.1109/ICBATS66542.2025.11258564
[25] Dip Bharatbhai Patel, Building scalable business intelligence systems in the cloud: A technical approach, International Journal of Science and Research Archive, 2021, 02(01), 212-215., Received on 03 February 2021; revised on 17 April 2021; accepted on 19 April 2021,DOI: https://doi.org/10.30574/ijsra.2021.2.1.0047
[26] Dip Bharatbhai Patel, University of North America, Virginia, USA, Home / Archives / Vol. 5 No. 02 (2025): Volume 05 Issue 02 / Articles, https://www.academicpublishers.org/journals/index.php/ijdsml/article/view/5958
[27] Sai Akshara Vemuri , KINEMATIC OPTIMISATION OF SOFT ROBOTIC MANIPULATORS FOR MEDICAL INTERVENTION, Vol. 54 No. 3 (2026): July-September 2026, System Protection and Control, ISSN-1674-3415, https://pspac.info/index.php/dlbh/article/view/450
[28] Sai Akshara Vemuri, https://www.periodicos.ulbra.org/index.php/acta/article/view/528
[29] Aziz, A., Chandusha, K., Naga Kiran, B., Prasad, M. L. M., Golla, K., Rajeswari, S., & Katta, S. K. R. (2026). Exploring Machine Learning Applications for Genomic Data Analysis in Personalized Medicine. International Journal of Drug Delivery Technology, 16(38s), 1166-1173. https://doi.org/10.25258/ijddt.16.38s.127,https://www.periodicos.ulbra.org/index.php/acta/article/view/687
[30] Mohammad Naffizuddin , Chaganti Ashok ,Sowjanya Gunukula , Rvbs Sarma , Bhavana Sujanamulk , Bharani Krishna Takkella , Chukka Ram Sunil ,Varri Sujana , Ashwini Kumar, Etiological Factors of Orofacial Clefts in a Hospital-Based Population: A Descriptive Analytical Study, Published: December 10, 2025,DOI: 10.7759/cureus.98868,DOI : https://doi.org/10.7759/cureus.98868
[31] Manoj Kumar Kagitha, Automated Security Compliance in Ci/Cd: A Natural Language Processing (NLP) Approach to Detecting Vulnerabilities in Infrastructure-As-Code (Terraform/Bicep), Publication year 2021, Volume 2021, Issue 12, https://computerfraudsecurity.com/index.php/journal/article/view/970
[32] Manoj Kumar Kagitha, Hybrid Quantum-Classical Resource Scheduling: A Proposed Architecture for Next-Generation Hyperscale Data Centers, Publication year 2024, Volume 2024, Issue 12, https://computerfraudsecurity.com/index.php/journal/article/view/969
[33] Manoj Kumar Kagitha, Self-Healing Cloud Infrastructure: A Hybrid Reinforcement Learning Framework for Predicting Microservice Failures in Azure Kubernetes Service (AKS),Publication year 2023, Volume 2023, Issue

12,https://computerfraudsecurity.com/index.php/journal/article/view/971
[34] Manoj Kumar Kagitha, GREENOPS IN THE CLOUD: AN AI-DRIVEN TELEMETRY ANALYSIS MODEL FOR REDUCING CARBON FOOTPRINT IN HIGH-AVAILABILITY CLUSTERS, Vol. 28 No. 6 (2020): JOCAAA, https://www.eudoxuspress.com/index.php/pub/article/view/5054
[35] Lakshmi Rahul Reddy Mareddy. (2026). Security and Performance Assessment of Encryption Techniques for Cloud Platforms. International Journal of Communication Networks and Information Security (IJCNIS), 18(1), 68–80. Retrieved from https://ijcnis.org/index.php/ijcnis/article/view/8735
[36] Lakshmi RahulReddy Mareddy, Context-Aware Learning Approaches for Improving Prediction Accuracy in Dynamic Systems, International Journal of Computational Mathematical Ideas (IJCMI), Vol.17 No.3(2025): IJCMI Vol17 NO3 2025 https://www.ijcmi.com/index.php/journal/article/view/15
[37] Lakshmi Rahul Reddy , Mareddy. (2026). Revisiting Autonomy in Artificial Intelligence through Agentic System Design. Journal of Computational Analysis and Applications (JoCAAA), 35(1), 707–713. https://eudoxuspress.com/index.php/pub/article/view/4780
[38] Mahendrakumar Kalal, Lean-Driven SAP Production Planning Optimization Using Machine Learning for Inventory and Throughput Efficiency, Date of Conference: 19-20 March 2026, Date Added to IEEE Xplore: 08 April 2026,ISBN Information:,ISSN Information: DOI:10.1109/ISDFS69419.2026.11459029, https://ieeexplore.ieee.org/abstract/document/11459029
[39] Mahendra kumar Kalal, SUPPLY CHAIN RESILIENCE THROUGH INTEGRATED SAP PRODUCTION PLANNING – PROCUREMENT SYNCHRONIZATION, Archives / Vol. 54 No. 1 (2026): January-March 2026 / Articles, https://pspac.info/index.php/dlbh/article/view/259
[40] Dr Manoj Goel (December 2018). THE ATTITUDE OF DIFFERENT STATES STUDENTS TOWARDS PHYSICAL EDUCATION; STUDYING IN PHYSICAL EDUCATION COLLEGES OF MAHARSHTRA STATE, https://ijeponline.com/index.php/journal/article/view/965/847
[41] Ilakiya Ulaganathan, Tagore Engg. College, Anna University, Chennai, India, ; FNU Antara; Santhosh Chitraju Gopal Varma; Zokir Mamadiyarov; Ziyadulla Turayev, Published in: 2025 International Conference on Sustainability, Innovation & Technology (ICSIT), Date of Conference: 22-23 August 2025, Date Added to IEEE Xplore: 19 December 2025, ISBN Information: https://ieeexplore.ieee.org/document/11294748/references#references
[42] Shreyas Kasture, Sudhanshu Maurya, Amit Kumar Sharma, Santhosh Chitraju Gopal Varma, Kashish Mirza, Firdous Sadaf Mohammad Ismail, DOI: http://doi.org/10.11591/ijai.v15.i2.pp1746-1761, IAES International Journal of Artificial Intelligence (IJ-AI), ISSN/e-ISSN 2089-4872/2252-8938 https://ijai.iaescore.com/index.php/IJAI/article/view/26932
[43] Shiv Shankar Dwivedi, Mr. Santosh Kumar Sharma, Future-Proofing Data Security with Quantum-Resistant Cryptography, Vol. 11 No. 3 (2025): IJCESEN, DOI: https://doi.org/10.22399/ijcesen.3447, https://ijcesen.com/index.php/ijcesen/article/view/3447
[44] Vijender Kalmotia, Enhancing statistical accuracy in gene perturbation studies, Volume 266, August 2026, 105819, https://doi.org/10.1016/j.biosystems.2026.105819, https://www.sciencedirect.com/science/article/pii/S0303264726001292
[45] Naveen Undrathi, Design and implementation of generative AI-driven data transformation pipelines using spring boot for scalable financial applications, eISSN: 2581-9615 || CODEN: WJARAI || Impact Factor 8.2 || CrossRef DOI, September 2026 (Volume 31, Issue 3), DOI url: https://doi.org/10.30574/wjarr.2026.30.3.1629,
[46] Naveen Undrathi, Agentic generative AI-enabled microservices framework

for intelligent financial data processing in java-based enterprise systems, September 2026 (Volume 31, Issue 3), eISSN: 2581-9615 || CODEN: WJARAI || Impact Factor 8.2 || https://doi.org/10.30574/wjarr.2026.30.3.1628

[47] Akshay Jain, Comparative Analysis of Linkage Methods in Hierarchical Clustering for Customer Segmentation, Conference paper, First Online: 07 October 2025, https://link.springer.com/chapter/10.1007/978-981-96-7289-9_44

[48] Ismail Hossain, Department of Computer Science, George Mason University, Fairfax, VA, USA, Mridul Banik, Department of Computer Science, Colorado State University, Fort Collins, CO, USA, https://ieeexplore.ieee.org/abstract/document/11446534 ,,DOI: 10.1109/FMLDS67896.2025.00106

[49] Ismail Hossain, Mridul Banik , Fahmid Al Farid , Jia Uddin , Hezerul bin Abdul Karim , Advancing Large Language Models for Low-Resource Languages: A Systematic Review of Pretraining, Adaptation, and Ethical Challenges, 10.32604/cmes.2026.075507, Date: 2026 https://www.sciencedirect.com/org/science/article/pii/S152614922600322X

[50] Ismail Hossain; Mridul Banik, Bridging Visual and Linguistic Intelligence for Chest X-rays: A Comprehensive Review of ViTs and LLM Synergies, Published in: 2025 IEEE International Conference on Future Machine Learning and Data Science (FMLDS) Date of Conference: 02-05 November 2025, Date Added to IEEE Xplore: 25 March 2026, ISBN Information: https://ieeexplore.ieee.org/abstract/document/11446534

[51] Rifat Sarker Aoyon1, Ismail Hossain2, M. Abdullah-Al-Wadud3, Jia Uddin4, A Secured and Continuously Developing Methodology for Breast Cancer Image Segmentation via U-Net Based Architecture and Distributed Data Training, https://www.sciopen.com/article/10.32604/cmes.2025.060917

[52] M. B. Vibha, M. Chandrika, Samitha Khaiyum & P. Rakshitha Kiran, Predicting sedentary behavior in adults using stacked LSTM modeling, published: 27 November 2024 Volume 16, pages 346–355 (2025), https://link.springer.com/article/10.1007/s13198-024-02622-2

[53] Bellary Chiterki Anil, Jayasimha Sondekoppa Rajkumar, T. L. Divya, Samitha khaiyum Rakshitha Kiran P., Balakrishnan Ramadoss, A Radiomics-based Framework for Liver Cancer Analysis using Explainable Artificial Intelligence (XAI) Methods, Vol. 15 No. 3 (2025): June, 2025, https://etasr.com/index.php/ETASR/article/view/10377

[54] Jayminkumar Patel, Sanjaykumar patel, Birju Patel, Rajan shah, Development and Validation of a Stability-Indicating Analytical Method for Quantitative Determination of Selected Anticancer Drugs in Pharmaceutical Dosage Form, Archives / Vol. 35 No. S2 (2026): Apr-jun 2026, https://welltestingjournal.com/index.php/WT/article/view/288

[55] Jayminkumar Patel, Sanjaykumar patel, Birju Patel, Rajan shah, Artificial Intelligence-Guided Optimization of Chromatographic Methods for Pharmaceutical Analytical Development, Archives / Vol. 35 No. S2 (2026): Apr-jun 2026, https://welltestingjournal.com/index.php/WT/article/view/294

[56] Sharoon Sharif, Inikpi Egbunu, Yujian Fu, and Zhigang Xiao. 2026. Hybrid Multi-Modal Learning for Extreme Long-Tail Fungi Classification. In 2026 ACM Southeast Conference (ACMSE 2026), April 23–25, 2026, Troy, AL, USA. ACM, New York, NY, USA, 38–45. https://doi.org/10.1145/3746467.3801502

[57] A. Chandrakumaran, "Innovative Strategies for Real-Time Analytics in Modern Data Warehousing Using Dockerized Ecosystems," 2025 International Conference on Advances in Next-Gen Computer Science

(ICANCS), Bangalore, India, 2025, pp. 1-6, doi: https://doi.org/10.1109/ICANCS65819.2025.11377462
[58] A. Chandrakumaran, "Enhancing SAP Predictive Maintenance and Industrial Analytics with LLM-Driven Automation," 2025 IEEE Madhya Pradesh Section Conference (MPCON), Jabalpur, India, 2025, pp. 238-244, doi: https://doi.org/10.1109/MPCON66082.2025.11256577.
[59] A. Chandrakumaran, "Enhancing SAP Cloud Optimization with Explainable AI: Cost Control, Workload Forecasting, and Performance Enhancement," 2025 3rd International Conference on Inventive Computing and Informatics (ICICI), Bangalore, India, 2025, pp. 456-463, doi: https://doi.org/10.1109/ICICI65870.2025.11069968.
[60] A. Chandrakumaran, "Innovative Strategies for Real-Time Analytics in Modern Data Warehousing Using Dockerized Ecosystems," 2025 International Conference on Advances in Next-Gen Computer Science (ICANCS), Bangalore, India, 2025, pp. 1-6, doi: https://doi.org/10.1109/ICANCS65819.2025.11377462.
[61] Ashwini Chandrakumara, year :2025, Leveraging Autonomous AI for Agile Growth: Transforming Project Management in Fast-Scaling Startups, Proceedings of the 2nd International Conference on Sustainable Business Practices and Innovative Models (ICSBPIM-2025), pages :730-743, issn :2352-5428, isbn :978-94-6463-872-1, Atlantis Press https://doi.org/10.2991/978-94-6463-872-1_46
[62] Anasuya Patil 1, Gurinderdeep Singh 2, Rajendra Dnyandeo Dighe 3, Dhruv Dev 4, Bhaveshkumar A Patel 5, Samatha Rudrangi 6, Gaurav Tiwari 7, Preparation, optimization, and evaluation of ligand-tethered atovaquone-proguanil-loaded nanoparticles for malaria treatment, 2025 Apr;36(6):711-742. doi: 10.1080/09205063.2024.2422704. Epub 2024 Nov 10., https://pubmed.ncbi.nlm.nih.gov/39522102/
[63] Kancharla Vijayalakshmi, Bethapudi Samuel Anand Andrews, Estimation of Cyproheptadine Hydrochloride and Tricholine Citrate Simultaneously in Syrup Dose by Applying Stability Indicating HPLC Methodology, December 31, 2020, Vol. 11, Issue 4, Pages 1813–1819, DOI: 10.26452/ijrps.v11iSPL4.4383, https://ijrps.com/v11si4/3683
[64] Vijayalakshmi Kancharla ,Kumar Abbaraju V D N, Andrews B S A, Nagaraju Dasari, Nagu Korupolu, Melphan as Anticancer and New Insights of Drug as Parenteral Dosage form against SARS-CoV-2 by Using Development and Validation , November 18, 2020, Vol. 11, Issue 1, Pages 1437–1446 DOI: 10.26452/ijrps.v11iSPL1.3686, https://ijrps.com/v11si1/2911
[65] Ramesh, D., A. Singh, and L. H. Channaiah. 2026. “ Evaluating the Effectiveness of Chewy Chocolate Chip Granola Bar Manufacturing Process on Salmonella Survival.” Food Science & Nutrition 14, no. 4: e71698. https://doi.org/10.1002/fsn3.71698.
[66] Inventor: Akhilesh RaiDilip Srinivas Krovi Sarthak Saptarshi Bandopadhyay Ravinder Singh Choudhary Richard LAWTON, 2025-02-28, Priority to US19/067,141,2025-11-14,Assigned to CITIBANK, N.A.,2025-12-30,Application granted,2025-12-30, Publication of US12511187B, https://patents.google.com/patent/US12511187B1/en?oq=US12511187
[67] Vivekanandan Govindan Ekambaram , REAL-TIME AUTOMATED DEFECT DETECTION IN INDUSTRIAL MANUFACTURING USING DEEP LEARNING AND TEXTURE ANALYSIS, Archives ,Vol. 54 No. 4 (2026): October-December 2026, Power System Protection and Control, ISSN-1674-3415,DOI: https://doi.org/10.46121/pspc.54.4.06, https://pspac.info/index.php/dlbh/article/view/545
[68] Vivekanandan Govindan Ekambaram, Yamuna Natarajan, AN OPTIMIZED HYBRID FRAMEWORK FOR REAL-TIME EDGE DETECTION AND FEATURE EXTRACTION IN AUTONOMOUS VEHICLE VISION SYSTEMS, Vol. 54 No. 4 (2026): October-December 2026 , Power System Protection and

Control, ISSN-1674-3415, DOI: https://doi.org/10.46121/pspc.54.4.05, https://pspac.info/index.php/dlbh/article/view/544
[69] John Smart Mgbakor, Behavioral Health Equity In High-Disparity Urban Communities: Public Health Impact, National Scalability, And Clinical Practice Implications, Pages1166 – 1177Cover Date, August 2026, Volume34, Issue, Special Issue 7- Fundamental Studies, Print ISSN-0965-7452, Electronic- ISSN-2396-8893, DOI: 10.1922/ejprd.v34i7s.1767, https://www.ejprd.org/article.php?article_id=1767&journal_id=171
[70] Nikhil Kumar Singh, Binay K. Singh, Piyush Kumar, Avinash Pandey, Sudhir Kumar, Sujit Kumar Bishi, A. Pattanayak, V. P. Bhadana, Sujay Rakshit & Kishor U. Tribhuvan, Integrative genomic and machine learning approaches reveal evolutionary signatures in the winged bean mitochondrial genome, Published: 28 November 2025,Volume 26, article number 3 (2026), https://link.springer.com/article/10.1186/s12870-025-07738-6
[71] GAURAV KUMAR SINHA, Year 2026 https://mathsapplication.com/article-view/?id=223
[72] GAURAV KUMAR SINHA, Year 2026 , https://mathsapplication.com/article-view/?id=224
[73] GAURAV KUMAR SINHA, Year 2026 , https://mathsapplication.com/article-view/?id=225
[74] Rasmi Ranjan Choudhury, Design and Implementation of a Tuned Rear Axle Bending Damper for Beat Boom Noise Mitigation in SUV Driveline Systems, 2026 International Conference, https://ieeexplore.ieee.org/document/11746870/authors#authors
[75] Rasmi Ranjan Choudhury, Comparative Performance Analysis of Single-Speed Versus Multi-Speed Transmissions in Battery Electric, 2026 International Conference, https://ieeexplore.ieee.org/document/11746704
[76] Nithesh Gudipuri1, Lalith Chandra Bandaru2, Mohammed Shakeer Bandrevu, SCALABLE DATA ENGINEERING FRAMEWORKS FOR REAL-TIME PROCESSING IN DISTRIBUTED SYSTEMS, Published:2026-10-18,Pages:444-453, https://actainnovations.com/index.php/pub/article/view/1900 ,doi: 10.62441/actainnovations.vi.1900
[77] Mohammed Shakeer Bandrevu; Pratham Patkar; Janardhana Naidu Kola, Trustworthy AI Models for Secure Data Processing Environments, Published in: 2026 International Conference on Natural Language Processing and Computer Vision (ICNPCV),Date of Conference: 17-18 July 2026, DOI: 10.1109/ICNPCV69434.2026.11708997, https://ieeexplore.ieee.org/document/11708997/authors#authors,
[78] Lalith Chandra Bandaru; Venkateswara Rao Batta; Uday Kumar Reddy Gangula, Hybrid AI and Data Analytics Frameworks for Cyber Resilience, 2026 International Conference on Natural Language Processing and Computer Vision (ICNPCV),Date of Conference: 17-18 July 2026,DOI: 10.1109/ICNPCV69434.2026.11708795, https://ieeexplore.ieee.org/document/11708795
[79] Lalith Chandra Bandaru; Karimulla Syed,Head of Access Management, S&P Global, Raleigh, NC, USA
[80] ; Nishanth Vepachedu; Rajmani Singh Oinam; Raja Vignesh Vaithiyalingam; Quazi Taif Sadat, Advanced Secure Knowledge Extraction Techniques for Large Language Models, 2026 International Conference on Emerging in Computing Technologies (INCECT),Date of Conference: 21-22 August 2026, DOI: 10.1109/INCECT72370.2026.11742788, https://ieeexplore.ieee.org/document/11742788
[81] Mohammed Shakeer Bandrevu, Independent Researcher, Aubrey, TX, USA, ; NagaSatyanarayana Raju Uppalapati; Ramesh Bellamkonda; Srikanth Kamatala, Llm-Augmented Cyber Threat Detection Using Federated Edge Intelligence and Data Mining Techniques, Published in: 2025 International

Conference on Computing Technologies & Data Communication (ICCTDC),Date of Conference: 04-05 July 2025, https://ieeexplore.ieee.org/document/11158153 , DOI: 10.1109/ICCTDC64446.2025.11158153

[82] Ankit Pathak, Leveraging AI for Better Data Quality and Insights, Volume 7 Issue 3 (2025), https://doi.org/10.32996/jcsts.2025.7.2.19, https://al-kindipublisher.com/index.php/jcsts/article/view/9345

[83] Uma Rajeswari Batchu, Bharati Reddi, Joshna Rani Surapaneni, Aruna Jangam, Sunil Misra, Anthony Addlagatta & Prakasham Reddy Shetty, Inhibition of Xanthine Oxidase by 1-O-Methyl Chrysophanol, A Hydroxyanthraquinone Isolated from Amycolatopsis Thermoflava ICTA 103, Published: 13 December 2025,Volume 66, pages 1049–1068 (2026) https://link.springer.com/article/10.1007/s12088-025-01534-4

[84] GAURAV KUMAR SINHA, Year 2025, https://visualcompublications.es/SAUC/article/view/6525

[85] GAURAV KUMAR SINHA, Year 2025, https://visualcompublications.es/SAUC/article/view/6526

[86] GAURAV KUMAR SINHA, Year 2025, https://visualcompublications.es/SAUC/article/view/6527

[87] D Benito Johnson 1, Gurinderdeep Singh 2, Deeksha Sharma 3, Venkatesan Natarajan 4, Knv Chenchu Lakshmi 5, Ram C Dhakar 6, Sadhana R Shahi 7, Suresh Velayutham 8, Ruchi Tiwari 9, Exploring Computational Advancements in ADME: Essential Insights for Drug Disposition, PMID: 39496312 DOI: 10.62958/j.cjap.2024.033, https://pubmed.ncbi.nlm.nih.gov/39496312/

[88] Mohammed Shakeer Bandrevu, Service Now Ebonding Methods, Implementation and Key features Published,30 Nov 2024, https://ijcttjournal.org/archives/ijctt-v72i11p125