Online Journal of Communication and Media Technologies, 2024, **14**(3), e202442
e-ISSN: 1986-3497

Research Article

The role of strategic online engagement and content curation in professional branding and career advancement on social media platforms

Pavel N. Ustin [1*]
0000-0003-3950-7434

Natalia N. Udina [2]
0000-0002-4945-7605

Elena V. Grib [3]
0000-0001-7191-1715

Roza L. Budkevich [4]
0000-0001-9221-090X

Andrey V. Korzhuev [5]
0000-0002-7454-038X

Nikolay N. Kosarenko [6]
0000-0002-5061-5551

[1] Kazan Federal University, Kazan, RUSSIA
[2] Peoples' Friendship University of Russia, Moscow, RUSSIA
[3] Moscow Aviation Institute (National Research University), Moscow, RUSSIA
[4] Almetyevsk State Technological University «Petroleum High School», Almetyevsk, RUSSIA
[5] I. M. Sechenov First Moscow State Medical University, Moscow, RUSSIA
[6] Plekhanov Russian University of Economics, Moscow, RUSSIA
[*] Corresponding author: pavust@mail.ru

Citation: Ustin, P. N., Udina, N. N., Grib, E. V., Budkevich, R. L., Korzhuev, A. V., & Kosarenko, N. N. (2024). The role of strategic online engagement and content curation in professional branding and career advancement on social media platforms. *Online Journal of Communication and Media Technologies, 14*(3), e202442. https://doi.org/10.30935/ojcmt/14772

ARTICLE INFO

Received: 29 Mar 2024

Accepted: 24 May 2024

ABSTRACT

This study investigates role of social media user engagement metrics in predicting career success likelihoods using supervised machine learning techniques. With platforms like LinkedIn and VKontakte becoming pivotal for networking and advancement, user statistics have emerged as potential indicators of professional capability. However, research questions metric reliability considering impression management tactics and biases. While prior studies examined limited activity features, this analysis adopts a robust CatBoost model to gauge career success prediction from multifaceted social data combinations. The study utilizes user profiles of over 17,000 on a major Russian platform. Individuals are categorized by an algorithm accounting for factors like salaries, experience, and employment status. User statistics spanning engagement, content sharing, popularity, and profile completeness provide model inputs. Following comparative evaluation, CatBoost achieved superior performance in classification accuracy, precision, recall and ROC AUC score. Analysis of SHapley Additive exPlanations values provides explanatory modeling insights into influential metrics, thresholds, and patterns. Results reveal subscribers, reposts and interest pages as highly impactful, suggesting that influence and content resonance predict success better than sheer visibility indicators like multimedia volumes. Findings also point to optimal engagement ranges beyond which career prediction gains diminish. Additionally, profile completeness and regular posting are positive to a limit, while likes to have negligible effects. The study contributes more holistic, data-driven visibility into effective social media conduct for career advancement. It advocates prioritizing network cultivation, tactical self-presentation, shareable narratives and reciprocal relationships over metrics gaming. Findings largely validate strategic communication theory around impression management and relationship-building.

Keywords: social media analytics, career advancement, personal branding, user engagement, VKontakte, professional networking

INTRODUCTION

With the integration of social media into professional networks and career development pathways, strategic online communication has become pivotal for nurturing opportunities and advancement (Duffy & Pooley, 2017; Koch et al., 2018). Platforms like LinkedIn, Twitter, and regionally popular VKontakte (VK) actively facilitate personal branding and access to career mobility resources (Donelan, 2016; van Dijck, 2013). Consequently, user engagement and popularity metrics on these sites have emerged as potential indicators of professional capability and success (Lardo et al., 2017; Mogaji, 2019).

However, research questions the reliability of metrics like followers, shares and post volumes as definitive representations of career achievement (Jacobson, 2020; Pennycook et al., 2021). Factors including tactical self-presentation, platform demographics and socio-cultural biases complicate interpreting online statistics as measures of professional competence (Brandtzaeg & Chaparro-Domínguez, 2020; Xu & Qian, 2023).

While prior studies have examined success prediction from limited activity features, a more holistic investigation of diverse user metrics can provide greater insight into effective career development communication norms on social media (Koltsova et al., 2017; Leite & de Baptista, 2022). Machine learning approaches that consider comprehensive indicator combinations offer data-driven assessments of impactful engagement strategies (Cerruto et al., 2022; Kostin & Shelukhin, 2016; Popov et al., 2021).

This study applies CatBoost algorithm, a state-of-the-art implementation of gradient boosting on decision trees developed by Yandex researchers. CatBoost has demonstrated superior performance compared to other traditional machine learning algorithms as well as advanced techniques like random forests and neural networks (Prokhorenkova et al., 2018; Ramaswami et al., 2022). Its key strength lies in efficiently handling categorical features with sophisticated encoding methods, automating various data preprocessing steps, and providing enhanced predictive power.

CatBoost's advantages include ordered target encoding for categorical features, which replaces labels with data-calculated representations to preserve the ordinal nature. It also employs oblivious trees, where each split is based only on the current feature value without considering previous splits, resulting in simpler trees that generalize better. Moreover, CatBoost introduces an innovative ordered boosting scheme that processes training examples in the order of strongest prediction errors, allowing earlier examples to influence the tree structure more (Asad et al., 2023).

Other notable features are the integration of ordered binning, random permutations during inferences to mitigate overfitting, and built-in over/underfitting detection mechanisms. CatBoost can automatically handle missing values, perform feature selection, and is robust to outliers (Ramaswami et al., 2022). With its ability to learn highly complex patterns and interactions, efficient memory management, and parallelization capabilities, CatBoost is particularly well-suited for large-scale datasets with mixed data types, high cardinality categories, and potential multicollinearity issues (Cakit & Dagdeviren, 2022).

This study applies CatBoost algorithm, a powerful gradient boosting framework that has demonstrated superior performance compared to other machine learning techniques by efficiently handling categorical features and automating various preprocessing steps (Prokhorenkova et al., 2018). CatBoost's advantages include ordered target encoding, oblivious trees, and an innovative ordered boosting scheme, making it particularly well-suited for complex datasets with mixed data types and potential multicollinearity (Asad et al., 2023).

VK platform serves as the investigation platform given user adaptations of its hybrid personal-professional environment for career aims in certain cultural settings (Donskoy et al., 2021; Orishev et al., 2020; Ruparel et al., 2020). Explanatory modeling with CatBoost reveals patterns in metric contributions to success predictions, informing strategic communication best practices.

The analysis makes three core contributions to research on social media metrics and career development strategies. First, it moves beyond evaluating isolated metrics to holistically examine success prediction from comprehensive combinations of diverse user data features using robust and advanced machine learning models like CatBoost. This contrasts with prior studies that relied on limited variable sets or less sophisticated techniques.

Second, by leveraging CatBoost's interpretability capabilities, the study provides nuanced visibility into the relative influence and thresholds of different metrics that contribute most substantially to success predictions. This granular insight challenges assumptions about which engagement factors definitively signal professional competence or achievement online.

Third, based on the influential metric patterns revealed, the analysis derives empirically-grounded, data-driven strategic communication principles to guide individuals in optimizing their professional social media presence for effective career development and mobility. These prescriptions account for the multifaceted realities of socially-rooted professional platforms.

More broadly, the investigation showcases the potential for deploying machine learning techniques to systematically evidence effective professional conduct norms and engagement strategies on emerging online networks that blend personal and career contexts. As these hybrid media spaces grow, data-driven insights mapping user behavior to tangible outcomes become increasingly vital for maximizing platform utilities.

LITERATURE REVIEW

Social Media's Evolution into Professional Networking Platforms

Social media has transitioned from its initial focus on personal interactions to a key tool in professional networking and career development. Initially designed for social engagement, platforms such as LinkedIn, Twitter, and VK have become essential for professionals looking to expand their networks, seek new opportunities, and advance their careers. This shift is supported by studies highlighting the platforms' evolution and their impact on professional practices (Orishev et al., 2020; Van Den Beemt et al., 2020; van Dijck, 2013).

The transformation from social to professional networking has been extensively analyzed, with early research indicating a trend, where users began to use their personal networks for professional gains (Kaplan & Haenlein, 2010). This trend evolved into a more strategic approach, where users curated their online profiles to reflect their professional identities, thus attracting more career opportunities (Boyd & Ellison, 2007; Brandtzaeg & Chaparro-Domínguez, 2020).

Comparative studies on platform effectiveness for professional networking show varied user behaviors and success rates. LinkedIn, known for its professional focus, is often highlighted for promoting behaviors that lead to career growth, contrasting with more socially-oriented platforms like VK, which may necessitate different strategies for professional networking (Koltsova et al., 2017; Mogaji, 2019).

Research specific to VK suggests its utility in regions with lower LinkedIn penetration, where users adapt its features for professional networking despite its social orientation (Cerruto et al., 2022). However, the blending of personal and professional content on such platforms requires nuanced management to maintain effectiveness (Donelan, 2016; Ruparel et al., 2020). The literature emphasizes the adaptive use of social media for professional ends, highlighting the platforms' iterative design changes and individual user innovations as key to their role in career development. While LinkedIn is a frontrunner in professional contexts, platforms like VK show potential in certain regions, underscoring the dynamic nature of social media's role in professional networking (Popov et al., 2021; van Dijck, 2013).

The progression of social media into professional networking is evident from the early 2000s with the emergence of LinkedIn, marking a significant shift from platforms focused on personal relationships to those facilitating professional connections. This evolution has been furthered by platforms like Facebook and Twitter, which have integrated business and job search functionalities alongside social features (Duffy & Pooley, 2017; Jacobson, 2020). The narrative of social media's expansion into professional domains is accompanied by changes in workplace communication norms, where personal and professional interactions increasingly intersect. This convergence is facilitated by technological advancements and platform features allowing for seamless transitions between personal and professional profiles, broadening access to professional networks and resources crucial for career advancement (Boyd & Ellison, 2007; Brodovskaya et al., 2019; Nikitkov & Sainty, 2014; Van Den Beemt et al., 2020).

Comparative analyses reveal distinct user behaviors across platforms like Twitter and LinkedIn, driven by their unique design features, which influence networking strategies and the effectiveness of professional interactions. Additionally, privacy controls and visibility settings play a significant role in moderating the impact of these networks, enabling users to manage their professional visibility and relationships strategically (Carpentier et al., 2019; Cerruto et al., 2022; El Ouirdi et al., 2015). Empirical studies on the career benefits of networking and personal branding on social media suggest a positive correlation between active engagement on platforms like LinkedIn and increased job opportunities. However, the challenge remains in establishing causality, with some evidence suggesting that social media can enhance job prospects and career advancement through strategic engagement and networking (Amadoru & Gamage, 2016; Buettner, 2017; Koch et al., 2018; Singh, 2023). The integration of social media into professional networking reflects its growing role in modern career development strategies. Ongoing research into platform features and user engagement strategies will continue to shed light on effective networking practices in the digital age.

User Engagement Metrics & Their Impact on Professional Opportunities

The integration of social media into professional networking has revolutionized personal branding and career advancement, making public metrics like likes, shares, and content frequency pivotal indicators of professional influence (Duffy & Pooley, 2017). User engagement on platforms such as LinkedIn and Twitter play a crucial role in showcasing expertise and leadership, directly impacting job attainment and career progression (Donelan, 2016; Mogaji, 2019).

Personal branding has evolved to encompass strategic content sharing, where engagement metrics such as likes, comments, and shares not only validate professional skills but also extend personal brand visibility (Jacobson, 2020). Regular content creation and sharing are associated with leadership perception and intellectual capital, enhancing professional connections and opportunities (Baima et al., 2020; Lardo et al., 2017).

Despite the apparent benefits, recent studies question the reliability of social media popularity metrics as true indicators of professional competence and achievement. Factors such as impression management tactics, the varying value of social media across professions, and socio-cultural biases complicate the interpretation of these metrics as definitive signs of career success (Marwick & Boyd, 2011; Pennycook et al., 2021). For instance, strategic behaviors like timing posts for maximum engagement and sharing emotionally charged content can artificially inflate popularity metrics, obscuring the relationship between online visibility and actual expertise (Bao, 2016; Xu & Qian, 2023). This discrepancy raises concerns about the legitimacy of using such metrics as reliable indicators of professional achievement.

While social media engagement metrics have become significant in personal branding and professional networking, their effectiveness and reliability as indicators of career success are nuanced and multifaceted. The interplay of strategic self-presentation, industry-specific values, and socio-cultural factors necessitates a more critical and nuanced approach to interpreting these metrics in the context of professional advancement. Further empirical research is essential to unravel the complex dynamics at play and to develop a more accurate understanding of how social media engagement translates into real-world career outcomes.

Strategic Communication for Career Advancement on Social Media

In an increasingly networked digital environment, strategic communication on platforms like LinkedIn and Twitter has become critical for career advancement through developing one's personal brand, expanding reach, and nurturing professional relationships (Duffy & Pooley, 2017; Mahmutova & Gerasimova, 2023; Sheer & Rice, 2017; Van Zoonen et al., 2016). This review analyzes recent research on effective practices for leveraging social media for career goals. Key areas examined include crafting an authentic personal brand, broadening network propagation, and converting online connections into opportunities through relationship-building.

Establishing a coherent personal brand constitutes foundational strategic communication for career advancement on social media (Carpentier et al., 2019; Jacobson, 2020; Tsaturyan & Matevosyan, 2022; Vartanov & Khvorostyanaya, 2023). This involves identifying and broadcasting one's unique expertise to establish professional credibility (Luwie & Pasaribu, 2021). For example, academics can selectively showcase research and commentary tailored for industry partners, policymakers or collaborators to expand their influence (Donelan, 2016). Regular posting that associates with leadership and intellectual capital also aids discoverability and career mobility (Brown et al., 2020; Kucharska, 2021; Luwie & Pasaribu, 2021). Overall, conveying specialized expertise and contribution builds legitimacy that can facilitate access to opportunities.

Once establishing a personal brand, leveraging one's networks for further visibility can amplify career development opportunities (Donelan, 2016; Duffy & Pooley, 2017). Tactical communication targeting high engagement can make brands "stickier" and expand real-world social capital (Lin et al., 2019). Techniques like emotional appeals increase content virality, though risk being manipulative (Marwick & Boyd, 2011). More sustainably, authentic conversation and mutual sharing build durable connections. Amplified online interaction metrics derive from and manifest real-world opportunity networks.

While visibility and broad networks create potential, strategic networking enables realizing career advancement from social media (Van Osch & Steinfield, 2018). These demands shifting communication from broadcasting to bilateral relationship cultivation based on reciprocity norms and appropriate self-disclosure (Chen et al., 2020; Leite & de Baptista, 2022; Marwick & Boyd, 2011). Such sincere relationship-building allows converting online networking into career mobility via referrals, endorsements and insider information. Sentiment analysis further enables gauging communication resonance to refine approaches (Umair et al., 2023).

Increasingly, online visibility and networks generate real-world career opportunities. Strategic communication for career advancement on social media involves developing personal brands, broadening reach through engagement techniques, and nurturing relationships to derive mobility benefits. Integrating data analytics also creates new avenues for evidence-based guidance and training around these core competencies. Additional research can clarify optimization approaches across diverse platforms and career contexts.

METHODOLOGY

The study adopts a quantitative methodology using supervised machine learning techniques to model career success prediction based on social media user data. Machine learning provides an effective approach for discerning complex patterns between multiple parameters that may not be tractable to traditional statistical methods (Osisanwo et al., 2017; Shetty et al., 2022). By training algorithms on labeled datasets, predictive insights can be derived on influential metrics and combinations that determine outcomes of interest (Enughwure & Ogbise, 2020; Tomasevic et al., 2020).

Data Collection Process & Sample

The data collection for this study was conducted using a specially developed information and analytical system for automated monitoring of personal pages on social networks. This system was designed to extract and analyze data from user profiles based on predefined criteria related to professional success. The development of this system utilized the Python programming language, leveraging VK API to access data from VK social network.

The system comprised two primary modules, as follows:

1. **Job applicant monitoring module:** This module focused on gathering information from specialized job search websites, identifying potential candidates based on their online profiles and submitted information.
2. **Social network monitoring module:** This component extended the search to VK social network, tracking personal pages of the applicants found on job search sites to gather additional data relevant to their professional and social engagement.

During the initial phase of testing, the system extracted a comprehensive range of data fields from the job search site databases, including personal information (e.g., name and age), professional details (e.g., desired position, salary expectations, work experience, and education), and other relevant data points (e.g., languages spoken, professional field, and region of job search). In addition to professional data, VK API was utilized to gather a range of social engagement metrics from the participants' VK profiles. These metrics included the number of friends and subscribers, the quantity of photos, videos, and audio recordings shared, the number of interesting pages or bookmarks, the completeness of profile information, and the volume of posts, reposts, and likes. These social metrics provided insights into the participants' online social behavior and engagement, contributing to a more nuanced understanding of their professional and personal profiles.

The research sample was refined through the system's testing phase, culminating in a total of 81,800 individuals. These subjects were primarily job applicants who had provided detailed information on specialized employment websites, making their profiles accessible to interested companies and recruiters. The demographic composition of the sample spanned a wide age range, from the onset of professional activity to retirement age, and included both male and female participants.

The applicants represented a diverse array of professional fields, aligning with a model that categorized 28 professional areas and 621 specific types of professional activities. This model was based on classifications used by HeadHunter, a leading job search platform in Russia, ensuring a comprehensive representation of various industries and job roles, such as those within the "automotive business" sector (e.g., Tinsmith, Auto Parts, and Car Washer).

The study's participants were further segmented into five categories based on their perceived professional success: unsuccess, rather unsuccessful, average, rather successful, and successful. This categorization was determined through an algorithm that considered multiple factors, including regional-adjusted wages, total work experience, tenure at the last job, current employment status, and foreign language proficiency.

Data Analysis

The primary research question of the study focused on predicting the likelihood of belonging to either a professional success or unsuccess based on various parameters derived from social network data. The data analyses were conducted using Python scripts, with the computational work performed on Google Colab.

Initially, from a total dataset of 81,800 records, data labeled as either successful or unsuccessful were extracted, resulting in a subset of 20,579 observations. Following this, a data cleaning process was undertaken to remove outliers and other anomalies, which reduced the dataset to 17,305 clean observations.

The cleaned dataset was then randomly split into training and testing sets, with 70% of the data allocated for training and the remaining 30% reserved for testing. This split was essential for validating the predictive performance of the models developed during the study.

Using the prepared training and testing sets, several machine learning algorithms were evaluated to identify the most effective model for classifying individuals into successful or unsuccessful groups. The algorithms considered in this analysis are listed in **Table 1**. The selection of the best-performing model was based on its accuracy, precision, recall, and F1-score on the testing set, alongside considerations of model interpretability and computational efficiency.

After identifying the optimal machine learning model, the study further investigated the importance of different features using SHapley Additive exPlanations (SHAP). SHAP values provide a powerful framework for interpreting machine learning models by quantifying the contribution of each feature to the prediction for individual observations, as well as on a more global scale. This analysis enabled the identification of key features that significantly influenced the likelihood of an individual being classified as successful or unsuccessful based on their social network profile. Features could include metrics such as the number of friends, subscribers, posts, likes, and the completeness of the profile, among others.

Beyond identifying important features, the study delved deeper into the effects of these variables on the predictions. This involved examining how changes in each feature value could alter the model's classification outcome for different observations. Such an analysis provides insights into the specific characteristics of social media profiles that are most predictive of professional success or failure.

Table 1. Machine learning algorithm & model selection criteria

Model	Accuracy	Precision score	Recall score	F1-score	ROC AUC score
XGBClassifier	0.816	0.822	0.920	0.868	0.767
LogisticRegression	0.777	0.775	0.931	0.846	0.705
KNeighborsClassifier	0.787	0.820	0.868	0.843	0.750
RandomForestClassifier	0.813	0.827	0.905	0.864	0.770
AdaBoostClassifier	0.802	0.814	0.906	0.857	0.753
DecisionTreeClassifier	0.733	0.797	0.798	0.797	0.702
MLPClassifier	0.777	0.807	0.870	0.837	0.734
SVC	0.804	0.803	0.929	0.862	0.745
SGDClassifier	0.756	0.760	0.921	0.833	0.679
GradientBoostingClassifier	0.810	0.819	0.914	0.864	0.762
LGBMClassifier	0.813	0.825	0.909	0.865	0.769
Catboost	0.820	0.830	0.915	0.870	0.776

RESULTS

Model Selection

Table 1 outlines a comparative analysis of various machine learning models based on several key performance metrics. These metrics gauge the effectiveness of each model in classification tasks. CatBoost emerges as a standout model with an accuracy of 0.82, indicating its superior ability to correctly predict outcomes 82% of the time, which is the highest among the models evaluated. Accuracy is a crucial metric as it reflects the overall correctness of the model, but it does not account for the balance between the classes. The precision score of CatBoost is 0.830, the highest on the list, showing its strength in generating relevant results over redundant ones. A high precision score is particularly important in scenarios, where the cost of false positives is high. Similarly, CatBoost's recall score is 0.915, which, while not the absolute highest, is competitively close. Recall is an essential metric when the consequences of false negatives are severe, as it measures the model's ability to capture all relevant instances. F1-score, which harmonizes the precision and recall into a single metric, is highest for CatBoost at 0.870. This suggests that CatBoost maintains an excellent balance between the precision and recall, an indication of its reliability and robustness in scenarios, where both false positives and false negatives carry a significant cost. Lastly, ROC_AUC score for CatBoost stands at 0.776, surpassing all other models. This metric is particularly telling of the model's capability to distinguish between the classes across different thresholds, which speaks to its generalization ability and diagnostic effectiveness.

CatBoost is the most favorable model across all the metrics considered, indicating consistent and reliable performance. Its superior accuracy, precision, balanced F1-score, and outstanding ROC_AUC score suggest that it would perform well in a variety of scenarios, making it a versatile and robust choice for classification tasks. The selection of CatBoost can be justified by its demonstrated ability to provide a reliable balance of error types and its expected performance on unseen data, which are critical factors for a practical and effective machine learning solution.

Importance of Independent Variables

SHAP summary plot (**Figure 1**) provided presents an insightful depiction of the influence various user engagement metrics have on the model's predictions regarding career success on VK social network. The most impactful feature, 'interesting pages', stands out prominently, implying that the number of interesting pages a user follows or interacts with has a substantial correlation with their professional achievements. This could be interpreted as a sign that engaging with content deemed valuable or relevant within their professional community is a strong indicator of career success.

Following 'interesting pages', 'subscribers' emerges as the second most significant feature. A higher count of subscribers might reflect a user's influence or reputation within the network, which in turn could correlate positively with their professional prospects. The 'reposts' metric also plays a considerable role, suggesting that users whose content is widely shared may enjoy greater visibility or prestige, potentially contributing to career advancement.

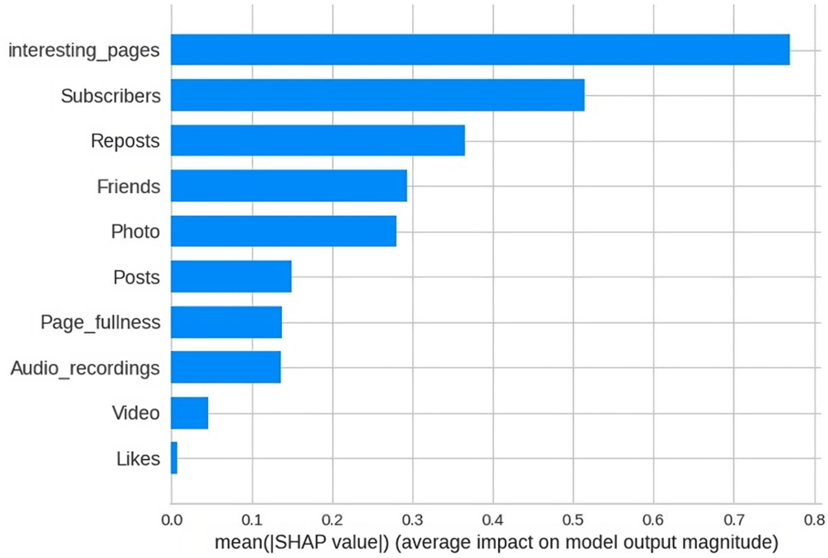

Figure 1. Importance of variables on model (Source: Authors, using study data from the SHAP library)

The 'friends' feature, while less impactful than 'reposts', still holds considerable weight. This indicates that a broad network of contacts might provide valuable opportunities, or alternatively, that successful individuals tend to cultivate larger networks. Similarly, the 'photo' metric underscores the role of active engagement, as the frequency of photo uploads could signal a user's level of interaction with their network, which may indirectly relate to their professional life.

'Posts' is another feature that the model identifies as influential. Regular posting activity could be a marker of a user's presence and visibility on the platform, which might parallel their professional engagement. Meanwhile, 'page fullness' has a moderate influence, suggesting that a comprehensive and detailed user profile may be somewhat beneficial for career success, possibly by enhancing personal branding or by providing a more complete picture of the user's professional persona.

Less influential in the model are 'audio recordings' and 'video', which suggest that these features, while part of a user's social media activity, do not have a strong direct correlation with career success according to the model's findings. Surprisingly, 'likes' rank as the least influential feature. Despite being a common metric for gauging content popularity, the number of likes a user receives is considered to have a negligible impact on career success predictions in this context. SHAP summary plot reveals that the model values the quality of user engagement and their influence within the social network higher than simple popularity metrics like likes or the number of multimedia uploads. It suggests a landscape, where professional success on VK is less about content consumption and more about meaningful interaction and influence within the user's network.

Two SHAP value bar charts (**Figure 2**) represent the impact of various features on the career success (chart a) and unsuccess (chart b) for users of a social networking app, VK. SHAP values explain the contribution of each feature to the model's prediction, where positive values indicate a push towards one outcome (success or unsuccess), and negative values indicate a push towards the opposite.

In the first chart, representing successful profiles, 'interesting pages' has the most positive impact with a SHAP value of +0.64, suggesting that engagement with content-rich pages is a strong predictor of career success. 'Subscribers' follows with a significant positive SHAP value of +0.36, reinforcing the idea that a higher subscriber count may enhance a user's professional image and potential opportunities. 'Posts' also contribute positively but to a lesser extent (+0.13), indicating that active posting can be a favorable factor, albeit less so than page interest and subscriber count. 'Page fullness' has a marginal positive effect (+0.04), hinting at the benefit of a complete profile, but its impact is minimal compared to the top features.

Conversely, 'reposts', 'friends', and 'photo' show negative SHAP values (-0.21, -0.17, and -0.13, respectively) in the success profile. Interestingly, this suggests that, for successful individuals, these features inversely correlate with career success, possibly implying that mere social activity may not be as critical as the quality of engagement.

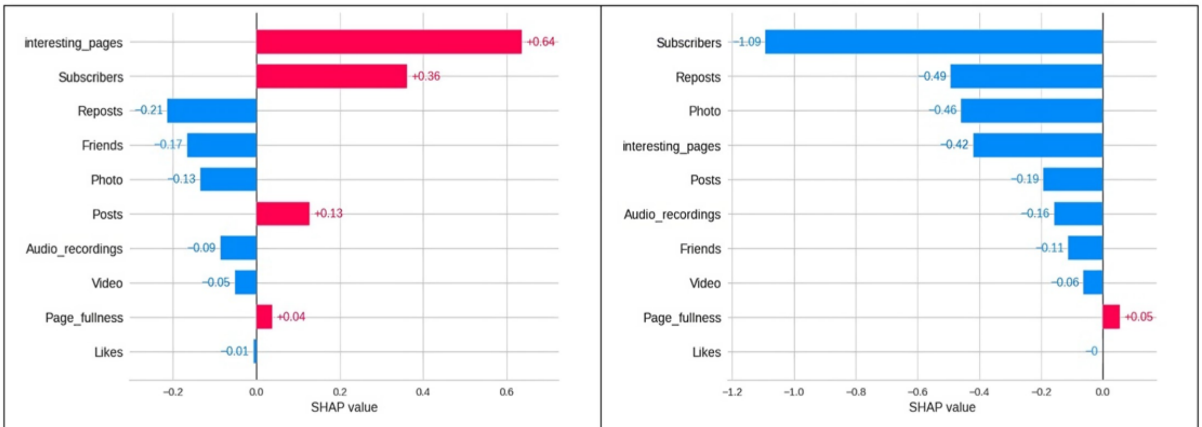

Figure 2. Features importance for each group (Source: Authors, using study data from the SHAP library)

The second chart, showcasing unsuccessful profiles, presents a contrasting view. Here, 'subscribers' has a highly negative SHAP value (-1.09), which indicates that having fewer subscribers is the most telling feature of a lack of career success. 'Reposts' and 'photo' also have negative impacts (-0.49 and -0.46), suggesting that less interaction in terms of content sharing and personal content contribution correlates with career stagnation. Notably, 'interesting pages' has a negative SHAP value (-0.42), in sharp contrast to its positive impact on successful profiles. This underlines that disengagement from meaningful content is a strong indicator of an unsuccessful profile.

Comparatively, between the two profiles, the most striking difference is with 'subscribers' and 'interesting pages' features. 'Subscribers' is a positive driver in the success chart and the most negative driver in the unsuccess chart, emphasizing the importance of how a vast subscriber base correlates with career outcomes. 'Interesting pages', while influential in both, switches from the most positive factor in successful profiles to a negative one in unsuccessful profiles, highlighting the significance of curated content engagement in professional success.

Other features like 'posts' and 'page fullness' show consistency between the two profiles, with 'posts' being slightly positive in both cases and 'page fullness' having a minimal impact. 'Reposts', 'friends', and 'photo' are negative for successful profiles but even more so for unsuccessful ones, reinforcing the notion that while social metrics are important, they might be less indicative of success than one's content curation and network influence. 'Audio recordings', 'video', and 'likes' to have minimal impact in both profiles, suggesting that these features are not significant predictors of career success or failure.

SHAP analysis reveals that for VK users, the factors that most strongly predict career success are not merely social activity metrics but rather indicators of influence and engagement with professionally valuable content. Conversely, the lack of such engagement and a smaller subscriber base are the most telling signs of an unsuccessful career profile according to the model.

Effect of Independent Variables

The scatter plot depicts the relationship between features a user has on the social network (VK) and the corresponding SHAP value for that feature in the context of the model predicting career success. SHAP values indicate the impact of a feature on the model's prediction for a particular outcome. In this plot, the horizontal axis represents feature user has, and the vertical axis represents SHAP value for that feature. Each point on the scatter plot corresponds to an individual prediction by the model. By examining the chart, it is determined in which range there is an increase or decrease.

Figure 3, the distribution of points suggests that as the number of subscribers increases, SHAP values also tend to increase, indicating a generally positive correlation between having more subscribers and the model's prediction of career success. Users with a higher number of subscribers appear to have positive SHAP values, meaning that this feature contributes to an increased prediction of career success. There are, however, some nuances, as follows:

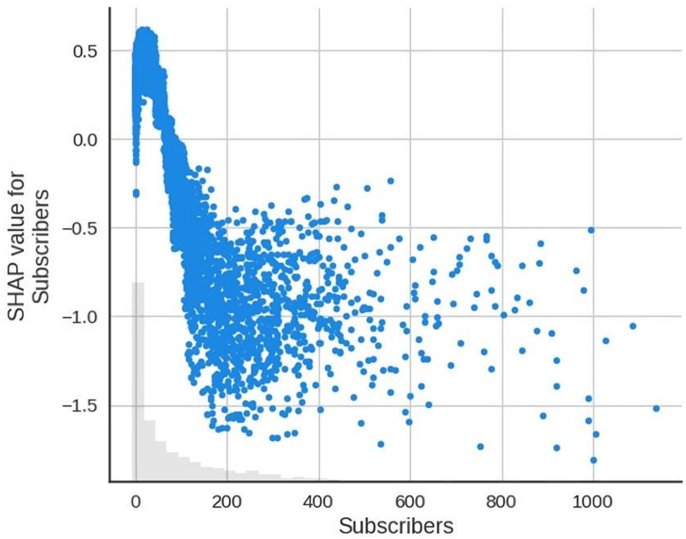

Figure 3. Scatter plot of subscriber & SHAP values (Source: Authors, using study data from the SHAP library)

For users with fewer subscribers, there is a wide distribution of SHAP values, ranging from slightly positive to highly negative. This implies that while having few subscribers can sometimes have a neutral or slightly positive effect on the model's prediction, it can also significantly decrease the predicted likelihood of career success. This wide distribution suggests that for users with a smaller subscriber count, other factors may be at play in determining their predicted career success.

For users with a moderate to high number of subscribers, SHAP values are predominantly positive but level off, indicating a diminishing marginal effect. In other words, after a certain point, gaining additional subscribers does not have as strong an impact on increasing the model's prediction of career success.

The histogram on the left side of the scatter plot shows the distribution of SHAP values for subscribers independently of the number of subscribers. The majority of SHAP values are concentrated around zero but with a long tail towards negative values. This distribution indicates that for a significant portion of the dataset, the number of subscribers is not a defining factor in predicting career success and that extremely low SHAP values (indicative of a negative impact on career success prediction) are less common.

As a results, the plot suggests that having a greater number of subscribers is generally associated with a higher prediction of career success by the model. However, the relationship is not linear, and there are diminishing returns to the positive impact of subscribers as their number increases. For those with fewer subscribers, the impact on the career success prediction is more varied, indicating that subscribers are an important but not exclusive factor in determining career success on VK according to the model.

Figure 4 indicates that as the number of interesting pages increases, SHAP values largely decrease. This suggests a negative correlation between the number of interesting pages and the model's prediction of career success, which is counterintuitive to what one might expect. Generally, one would assume that engagement with a higher number of interesting pages might indicate a broader interest or more significant professional engagement, which could positively influence career success.

The concentration of dots at the lower end of the 'interesting pages' scale, with higher SHAP values, might indicate that a moderate level of engagement with interesting pages contributes positively to career success predictions. However, as the number of interesting pages increases, their contribution to career success predictions becomes negative. This could imply that beyond a certain point, excessive diversification in page interests may be perceived by the model as less indicative of a focused professional profile, which could negatively impact career success predictions.

The histogram on the left side of the plot shows the distribution of SHAP values for 'interesting pages', independent of the actual number of pages. It demonstrates that most SHAP values are slightly negative, suggesting that for the majority of the dataset, higher engagement with interesting pages does not correspond to a prediction of higher career success.

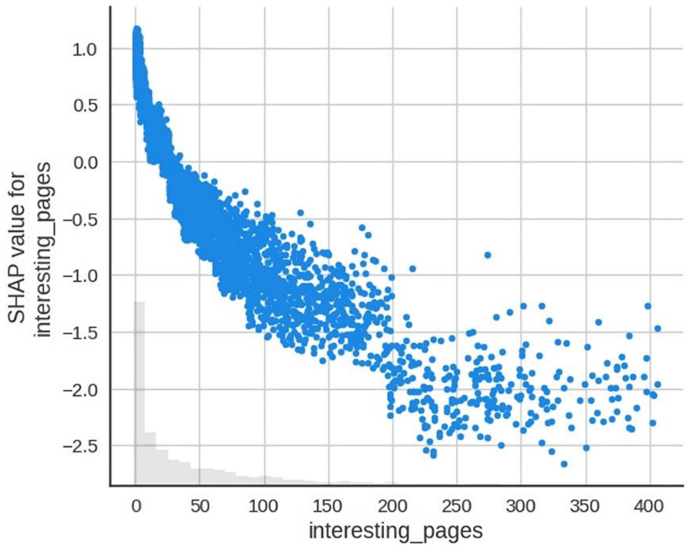

Figure 4. Scatter plot of interesting pages & SHAP values (Source: Authors, using study data from the SHAP library)

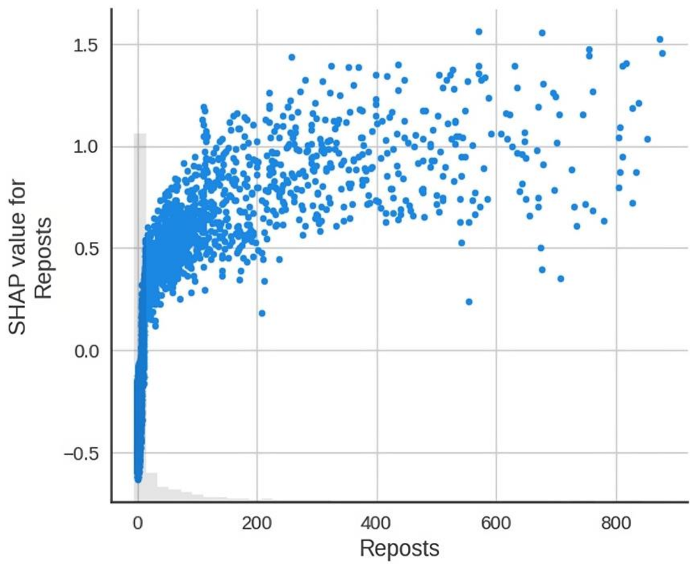

Figure 5. Scatter plot of reposts & SHAP values (Source: Authors, using study data from the SHAP library)

This scatter plot challenges the assumption that more engagement (in terms of the number of interesting pages a user interacts with) necessarily translates to a prediction of higher career success. Instead, it presents a more nuanced picture, where a balanced level of engagement with interesting pages may be optimal, according to the model's assessment.

For a large number of users with few reposts (close to zero), SHAP values span a range from negative to slightly positive. This indicates that having few reposts can either slightly decrease or increase the model's prediction of career success. The variation in SHAP values suggests that other factors might be influencing the prediction for users with a low number of reposts (**Figure 5**).

As the number of reposts increases, SHAP values tend to cluster around the zero to slightly positive range. This pattern implies that, according to the model, as users have more content being reposted, the impact on the predicted career success is generally positive but not strongly so. There is a saturation point beyond which additional reposts do not significantly increase the prediction of career success.

There are a small number of users with a very high number of reposts, and for these users, SHAP values are mostly positive, although they do not appear to increase proportionally with the number of reposts. This

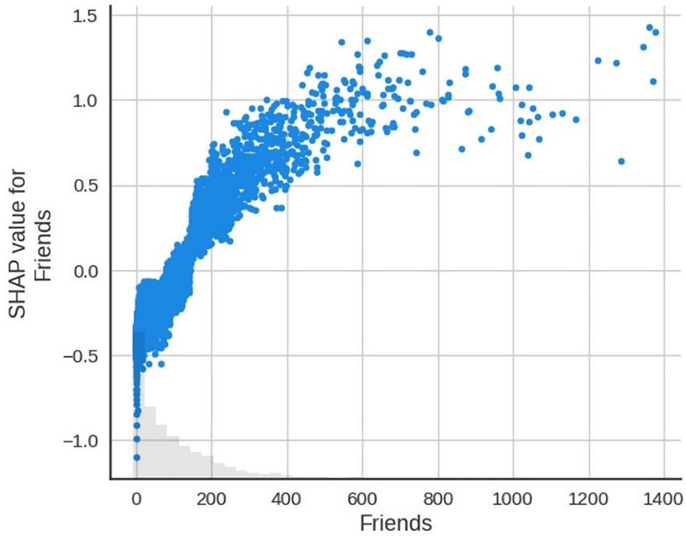

Figure 6. Scatter plot of friends & SHAP values (Source: Authors, using study data from the SHAP library)

suggests a diminishing return effect, where, beyond a certain number of reposts, the positive impact on career success predictions does not continue to grow.

The histogram on the left side of the scatter plot shows the distribution of SHAP values for reposts, independent of the number of reposts. This distribution is heavily skewed towards the lower end, indicating that for most of the dataset, having more reposts does not drastically change the prediction of career success.

As a result, SHAP scatter plot for reposts illustrates that while having more reposts generally contributes positively to the model's prediction of career success on VK, the effect plateaus after a certain point. This could reflect the model's inference that while user engagement, as measured by reposts, is beneficial to career success, there is a limit to its impact, and other factors likely come into play.

For users with a relatively low number of friends, SHAP values range from negative to neutral, suggesting that within this range, having fewer friends may either decrease the model's prediction of career success or have a negligible impact (**Figure 6**).

As the number of friends increases, we see an ascent in SHAP values, which then plateau. This indicates a positive correlation between the number of friends and career success predictions up to a certain point. After this point, additional friends do not appear to significantly alter the predicted success. The plateau of SHAP values in the higher friend counts implies a diminishing return effect, similar to what we observed with the number of reposts. This means that while a larger network may initially contribute positively to career success predictions, there's a threshold beyond which the size of the network does not continue to have a strong positive impact. There's a considerable concentration of dots at the lower end of the friend count, which transitions into a sparse distribution as the friend count increases. This spread suggests variability in the model's interpretation of the importance of having a vast number of friends.

The histogram on the left side of the plot illustrates the distribution of SHAP values for the number of friends. It shows a high concentration of SHAP values around zero, indicating that for a substantial portion of the dataset, the number of friends is not a decisive factor in the model's prediction of career success. The concentration of negative SHAP values for lower friend counts suggests that having very few friends is more frequently associated with a decrease in career success predictions. As a results, SHAP scatter plot for the 'friends' feature suggests that having a moderate number of friends on VK is associated with a higher prediction of career success, but after reaching a certain network size, the marginal benefit decreases. This might imply that the model finds a balanced network size to be optimal for career success beyond which the value of each additional friend diminishes in terms of impact on the career success prediction.

A substantial number of users with a low photo count have SHAP values clustered near zero, but with a tail extending into negative values (**Figure 7**). This suggests that for users with fewer photos, this feature may either slightly negatively impact the model's prediction for career success or have little to no impact. As the

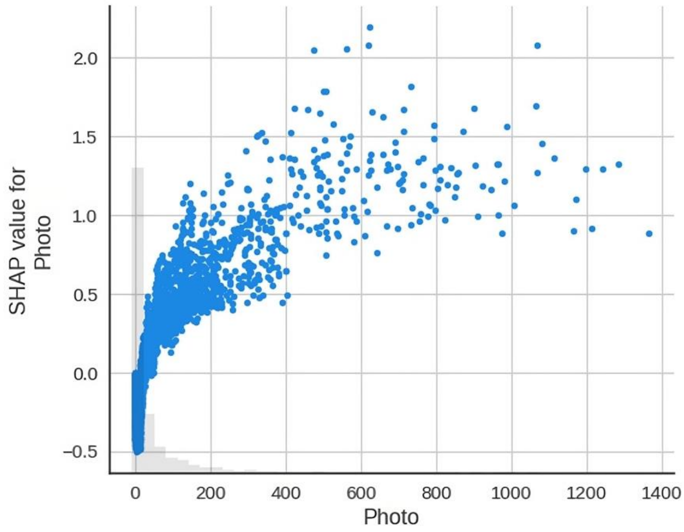

Figure 7. Scatter plot of photo & SHAP values (Source: Authors, using study data from the SHAP library)

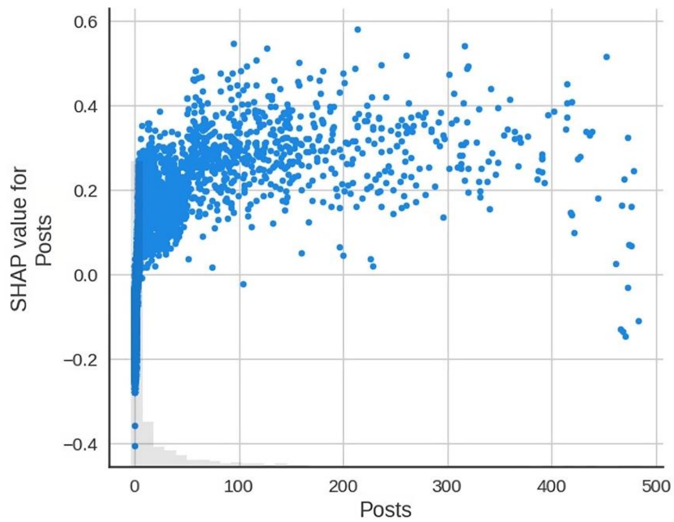

Figure 8. Scatter plot of post & SHAP values (Source: Authors, using study data from the SHAP library)

number of photos increases, SHAP values start to spread out, with a trend toward positive values. This indicates that a higher photo count may contribute positively to the model's prediction of career success.

However, the distribution of SHAP values is wide, indicating variability in how much the photo count impacts the predicted success. There does not appear to be a clear linear relationship between the number of photos and SHAP values, as seen by the wide dispersion of points. This may imply that while posting photos is generally seen as a positive action by the model, other factors may influence the extent of its impact on career success. For users with a very high number of photos, SHAP values are predominantly positive, yet they plateau, indicating diminishing returns. This pattern suggests that beyond a certain point, additional photos do not contribute significantly to higher predictions of career success. The histogram on the left side of the plot shows the distribution of SHAP values for 'photo'. There is a heavy concentration of SHAP values at the lower end, indicating that for most of the dataset, having a low to moderate number of photos does not substantially change the prediction of career success. The number of photos only starts to have a more positive impact on career success predictions as the count increases beyond the average range. The scatter plot implies that active photo posting on VK may be associated with higher predictions of career success, with the caveat that beyond a certain level of activity, the benefit to the prediction does not increase significantly. This suggests that the model considers photo posting as one indicator of a user's engagement or visibility, which could be related to career success, but with a nuanced interpretation that acknowledges the limits of this effect.

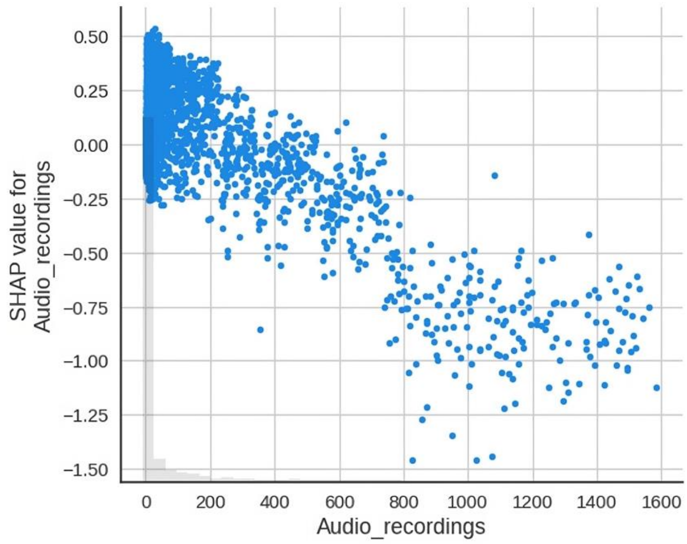

Figure 9. Scatter plot of audio recording & SHAP values (Source: Authors, using study data from the SHAP library)

For users with a small number of posts, there's a noticeable cluster of SHAP values at or slightly below zero (**Figure 8**).

This suggests that not posting or posting very little has a neutral or slightly negative effect on the model's prediction of career success. The cluster's presence near the axis indicates that for these users, 'posts' is not a strong predictor of career success. As the number of posts increases, SHAP values appear to increase as well but then plateau. This indicates that while an increase in posting frequency can positively influence the model's prediction of career success, it does so up to a certain point. Beyond this point, the contribution of the number of posts to the predicted career success levels off. The distribution of SHAP values for users with a higher number of posts does not show a strong positive or negative trend. Instead, it demonstrates a relatively stable influence on the model's predictions, with SHAP values mostly scattered above zero. There are no extreme positive SHAP values, even as the number of posts grows large. This suggests that the model does not consider a high posting frequency to be a very strong indicator of career success on its own. The histogram to the left indicates the distribution of SHAP values across all users. Most values are concentrated around zero, particularly for users with fewer posts, reflecting a neutral impact on the model's output. As the number of posts grows, SHAP values tend to be positive, but the increase is not proportional to the number of posts. The scatter plot illustrates that while posting more frequently on VK may have a positive impact on a user's predicted career success, the effect is moderate and plateaus beyond a certain level of activity. This indicates that while the model recognizes the importance of user engagement through posting, it also suggests there may be an optimal range of activity that correlates with career success after which the incremental benefit of additional posts diminishes.

A large cluster of data points near the origin suggests that for users with a low number of audio recordings, this feature does not have a significant impact on the model's prediction of career success (**Figure 9**). SHAP values are very close to zero for a wide range of users with few audio recordings. As the number of audio recordings increases, SHAP values become more spread out and generally trend downwards. This trend indicates that a higher count of audio recordings may be negatively associated with the model's predictions of career success. There is a noticeable concentration of negative SHAP values as the number of audio recordings increases, suggesting that the model may interpret a high volume of audio content as a less favorable indicator for career success. However, the negative impact is not linearly proportional to the number of recordings; rather, it appears to level off after a certain point. The distribution of SHAP values does not display a significant positive influence for any count of audio recordings. Unlike other features such as posts or photos, where more activity could correlate with increased career success, audio recordings do not show a similar pattern. The histogram to the left shows a heavy concentration of SHAP values at the lower end for users with fewer audio recordings, indicating that the 'audio recordings' feature has a neutral to negative impact on career success predictions for the majority of the dataset. The scatter plot suggests that

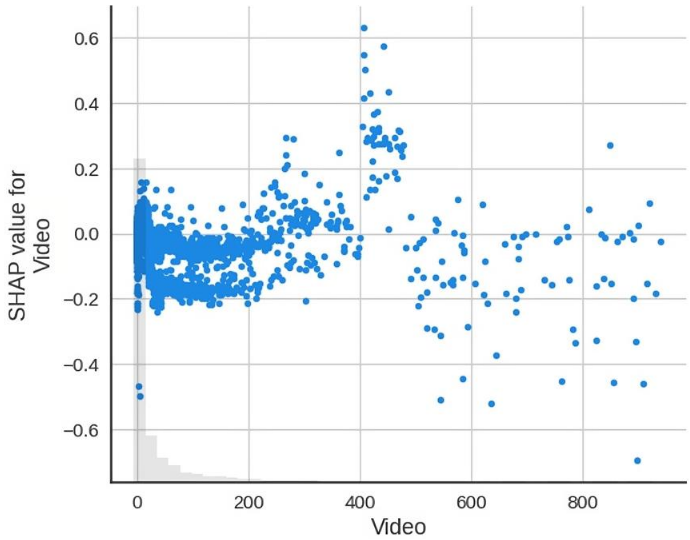

Figure 10. Scatter plot of video & SHAP values (Source: Authors, using study data from the SHAP library)

the number of audio recordings a user has on VK is generally not positively associated with the model's predictions of career success. For users with a moderate to high number of audio recordings, this feature could even be a slight detriment to the predicted outcome. The model likely interprets engagement through audio content differently than other types of engagement, such as posting text or photos, when assessing factors that contribute to career success.

For users with a small number of videos, there's a dense cluster of SHAP values near zero, with a tail extending slightly into negative values (**Figure 10**). This suggests that for users with few videos, this feature has a neutral to marginally negative impact on the model's prediction of career success. It indicates that at lower levels, the 'video' feature is not a strong determinant in the model's assessment of career success. As the number of videos increases, SHAP values spread out with a slight upward trend, indicating that having more videos could potentially contribute positively to career success predictions. However, the spread of SHAP values also suggests variability in how much the video count impacts the predicted success across different users. There is no clear indication of a strong positive correlation between the number of videos and SHAP values. Despite some higher values being present for users with more videos, the overall pattern does not strongly suggest that increasing video content leads to a higher prediction of career success. The distribution of SHAP values does not show a significant downward trend even as the number of videos grows, suggesting that while the 'video' feature may not be a strong positive predictor of career success, it is also not a strong negative predictor. The histogram on the left side of the scatter plot shows the distribution of SHAP values for the 'video' feature across all users. The majority of values are concentrated around zero, especially for those with fewer videos, indicating that for most of the dataset, the 'video' feature does not have a decisive impact on career success predictions. The scatter plot indicates that the number of videos a user posts on VK might have a slightly positive impact on the model's prediction of career success, but this effect is not definitive or strong. The feature exhibits a varied influence on the prediction, with no significant positive or negative trend as the video count increases. This suggests that while the model recognizes the presence of video content as a component of a user's social media activity, it is likely one of many factors, and not a dominant one, in the determination of career success according to this model's assessment.

SHAP values are predominantly centered around zero, especially for mid-range values of 'page fullness' (**Figure 11**). This central clustering suggests that for a typical level of profile completeness, the impact on the model's prediction of career success is neutral. There is a noticeable spread of SHAP values at lower levels of 'page fullness', with some negative values, which could imply that incomplete profile pages may sometimes contribute to a lower prediction of career success. As 'page fullness' increases, SHAP values seem to flatten out, indicating that beyond a certain threshold of completeness, adding more information to the page does not significantly impact the model's prediction. This suggests that the model might value a moderate level of profile completeness, but there are diminishing returns after the profile meets a certain standard of fullness. There are no strong positive SHAP values for any level of 'page fullness', which would indicate that a very full

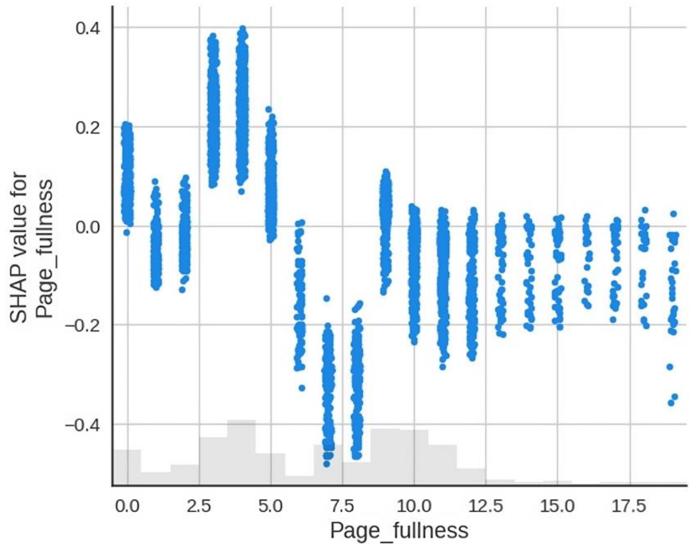

Figure 11. Scatter plot of page fullness & SHAP values (Source: Authors, using study data from the SHAP library)

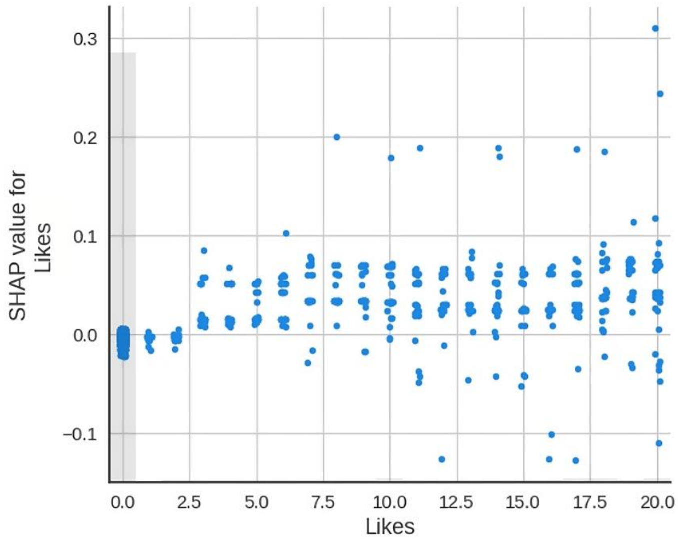

Figure 12. Scatter plot of likes & SHAP values (Source: Authors, using study data from the SHAP library)

page is not necessarily seen as a strong predictor of career success by the model. The histogram on the left side of the scatter plot shows the frequency of SHAP values across all users.

The highest concentration of SHAP values is around zero, reinforcing the point that 'page fullness' has a neutral effect on the model's output for most users. The scatter plot implies that while having a certain level of completeness on one's VK profile page ('page fullness') might be associated with the model's predictions of career success, there is no strong correlation that suggests increasing completeness continues to significantly impact those predictions. The model seems to consider a moderately full page as potentially beneficial, but after reaching a certain level of completeness, additional information does not considerably change the prediction of career success.

SHAP values for 'likes' are tightly clustered around zero, especially as the number of likes increases (**Figure 12**). This tight clustering indicates that the number of likes has a relatively small impact on the model's predictions of career success for VK users. There is a lack of any strong positive or negative trend in SHAP values with respect to the number of likes. This implies that receiving more likes does not necessarily translate to a higher or lower prediction of career success by the model. The range of SHAP values for 'likes' is quite narrow, staying within approximately -0.10 to +0.30. Such a narrow range suggests that 'likes' are not a significant driver of the model's predictions, having only a marginal effect, whether positive or negative. Even users with a higher count of likes do not exhibit large SHAP values, further reinforcing the idea that likes are not a major factor in the model's assessment of a user's career success. The histogram on the left side of the plot shows the distribution of SHAP values for the 'likes' feature across all users.

Most of SHAP values are concentrated near zero, regardless of the number of likes, underscoring the minimal influence this feature has on the model's predictions. The scatter plot suggests that while likes are a visible and commonly-tracked metric on social media platforms, they do not significantly influence the model's predictions of career success for VK users. According to this model, likes are likely considered less indicative of professional achievement or potential than other factors.

DISCUSSION

The findings from the analysis of user engagement metrics on VK using CatBoost machine learning model offer profound insights into the nexus between social media behavior and professional success. These insights not only corroborate existing scholarly discourse but also extend our understanding of strategic communication and personal branding within digital ecosystems. Here, we delve deeper into the implications of these findings, particularly through the lens of communication theories and practices.

The significance of cultivating a dedicated subscriber base as a marker of career success underscores the pivotal role of personal branding in the digital age (Blaer et al., 2020; Labrecque et al., 2011). This aligns with the concept of 'micro-celebrity' status, where individuals adopt strategies akin to celebrities to build a following and engage with an audience on social media (Khamis et al., 2017). The strategic self-presentation and content curation that led to a growing subscriber base can be viewed through the prism of Goffman's self-presentation theory, emphasizing the performative nature of social interactions, extended into the digital realm (Merunková & Šlerka, 2019). Creating an influential presence by cultivating subscribers is validated as an impactful success metric (Cerruto et al., 2022; Mogaji, 2019), affirming the importance of developing a strong personal brand and reputation that garners a dedicated audience. Strategic self-presentation that signals expertise and leadership can contribute to assembling a subscriber base indicative of professional capability (Koltsova et al., 2017; Ruparel et al., 2020).

The concept of reposts is pivotal in understanding how digital content can effectively engage audiences by aligning with their core values and interests. According to the uses and gratifications theory, as discussed by Ruparel et al. (2020), individuals are not passive consumers of media but active selectors who choose media sources that fulfill their specific needs and desires, a concept further explored by Xie and Liu (2022). This theoretical framework suggests that when content is frequently shared or reposted, it signifies that it has successfully met the audience's expectations and desires, whether these are informational, entertainment-based, or social. Moreover, the frequency of reposts serves as a barometer for content's resonance with an audience's cultural and social values. As Chen et al. (2020) and Marwick and Boyd (2011) articulate, content that aligns well with community norms and values tends to gain greater visibility and create more engagement opportunities. This phenomenon underscores the importance of understanding the socio-cultural dynamics that influence audience behaviors online.

In crafting content that is likely to be shared widely, it is essential to weave narratives that not only tap into the audience's emotional needs but also adhere to and reflect conversational norms and cultural subtleties, as suggested by Bao (2016) and Xu and Qian (2023). These narratives should skillfully incorporate elements that trigger emotional responses and foster a sense of community and belonging among the audience, thereby enhancing the likelihood of reposts and further dissemination. This strategic approach to content creation highlights the interplay between effective communication, cultural understanding, and emotional engagement in the digital media landscape.

The negative correlation between engagement with a wide array of "interesting pages" and professional success challenges the conventional wisdom that broader visibility invariably leads to better career outcomes. This suggests that strategic communication involves prioritizing focused, reputation-aligning interests over diverse visibility (Carpentier et al., 2019; Popov et al., 2021). Spreading engagement across too many domains diffuses perceptions of specialized expertise that underpin leadership legitimacy. Also, this finding can be interpreted through the lens of the communication theory of identity, which posits that identity is constructed across four frames: personal, enactment, relational, and communal (Kuiper, 2021). Excessive diversification may blur these identity boundaries, leading to a less coherent personal brand. A focused approach, on the other hand, helps in maintaining a clear and consistent identity that is more likely to be perceived as authentic and trustworthy.

Similarly, while a broad friend network initially correlates positively with success, connections must cultivate meaningful reciprocity and insider access to drive career mobility (Amadoru & Gamage, 2016; Macy, 2023; Singh, 2023; Van Osch & Steinfield, 2018). Communication that strategically balances appropriate self-disclosure and bilateral relationship-building is key to materializing opportunities (Leite & de Baptista, 2022; Sheer & Rice, 2017). Also, this result underscores the complexity of social capital in digital spaces (Nikitkov & Sainty, 2014). While a large network provides visibility, the depth and quality of connections–characterized by meaningful reciprocity and insider access–are crucial for leveraging this network for career mobility. This aligns with Putnam's distinction between bridging (inclusive) and bonding (exclusive) social capital, where the latter, characterized by stronger, more meaningful connections, is more influential in facilitating career advancement (Putnam, 2000).

The findings challenge the overemphasis on volume metrics (likes, video and audio posts) as indicators of professional success, inviting a reassessment of what constitutes valuable engagement in professional contexts. This reflects the media richness theory, which suggests that the effectiveness of communication is contingent on the medium's capacity to convey rich information (Wang et al., 2022). In professional settings, content that conveys expertise and substantive insight may be valued over sheer volume, highlighting the need for strategic content that aligns with professional goals. This comports with evidence questioning the reliability of impressions management tactics in signaling professional capability (Jacobson, 2020; Pennycook et al., 2021). It also underscores the variance of metric value across industries (Koltsova et al., 2017).

These insights illustrate that effective social media communication for career advancement transcends mere visibility and encompasses strategic networking, nuanced personal branding, and the crafting of narratives that resonate with specific communities. The findings advocate for a strategic approach to social media engagement that prioritizes quality and relevance over quantity, emphasizing the construction of a coherent personal brand and the cultivation of meaningful, reciprocal connections.

These insights illustrate that effective social media communication for career advancement transcends mere visibility and encompasses strategic networking, nuanced personal branding, and the crafting of narratives that resonate with specific communities. The findings advocate for a strategic approach to social media engagement that prioritizes quality and relevance over quantity, emphasizing the construction of a coherent personal brand and the cultivation of meaningful, reciprocal connections.

Building on these insights, future research could explore the synergistic effects of various communication strategies and their alignment with career goals across different industries. Investigating the interplay between personal and professional identities on social media and its impact on career success could provide further depth to our understanding of digital personal branding. Additionally, cross-platform studies could illuminate how platform-specific affordances and user demographics influence the relationship between social media behavior and professional outcomes, offering a more holistic view of strategic communication in the digital age. By weaving these nuanced findings with established communication theories, we gain a richer understanding of the dynamics at play in leveraging social media for professional success, offering a roadmap for individuals and professionals navigating the digital landscape.

CONCLUSIONS

The comprehensive analysis of various machine learning models and their respective performance metrics reveals that CatBoost model stands out with superior accuracy, precision, balanced F1-score, and ROC AUC score. This model's ability to provide consistent and reliable performance across different evaluation metrics makes it an exemplary choice for classification tasks related to predicting career success on social media platforms like VK.

From the perspective of communication, the significance of user engagement metrics on VK is intricately linked with the indicators of career success as interpreted by CatBoost model. The analysis of SHAP values offers nuanced insights into how different aspects of a user's online presence contribute to their professional image and potential career advancement. Notably, the number of interesting pages a user follows or interacts with is the most impactful feature, highlighting the value of engaging with professionally relevant content over merely popular content.

Subscribers and reposts also emerge as significant features, implying that a user's influence and the shareability of their content are strong indicators of career success. The number of friends and the frequency of photo uploads illustrate the importance of an active digital presence and a broad network, albeit with diminishing returns beyond a certain point. Regular posting activity, while influential, reaches a plateau, suggesting that an optimal range of activity correlates with career success beyond which additional engagement does not significantly enhance the model's success prediction. Surprisingly, common metrics like audio recordings, video content, and likes are not strong predictors of career success. This points to a potential discrepancy between perceived popularity and professional achievements as per the model's findings. In terms of communication, this analysis underscores the importance of quality interactions and a well-curated online presence. It challenges the conventional wisdom that more content and higher engagement always correlate with better professional outcomes. Instead, it proposes a more strategic approach to social media usage, where the focus on relevant content and meaningful connections aligns more closely with indicators of career success. Ultimately, this investigation into machine learning models not only guides the selection of robust predictive tools but also provides valuable insights for users seeking to optimize their social media behavior for career growth. It emphasizes the strategic importance of building a substantial yet relevant network, sharing content that resonates with one's professional community, and cultivating a comprehensive online profile that reflects one's personal brand and professional expertise.

The findings from this study highlight several best practices for leveraging social media platforms like VK for career advancement. Firstly, cultivate a dedicated base of followers/subscribers through strategic self-presentation and content sharing that signals thought leadership. This underscores the importance of developing a strong personal brand that garners a loyal audience. Secondly, prioritize engaging with professionally relevant content and influential figures over diverse interests. Excessive diversification can dilute perceptions of specialized expertise that drive leadership legitimacy. Maintaining a focused approach aids in coherent personal branding. Thirdly, craft narratives and messaging that resonate with the target community's values and tap into emotional appeals for increased virality. This fulfills audience needs for entertainment, belonging and information sharing. Fourthly, convert broad networks into meaningful reciprocal relationships characterized by appropriate self-disclosure and insider access opportunities. The depth of connections drives career mobility more than sheer network size. Finally, individuals should reorient metrics of success away from volume of generic content consumption towards indicators of influence and resonance within key professional communities.

While this study provides valuable insights, there are certain limitations. The analysis relies exclusively on VK user data. Additional research should investigate how findings translate across platforms with different user demographics and platform affordances. Professional success definition and contributing metrics likely vary across cultural settings and industry contexts. More focused situational analysis can reveal nuances around success metrics. Causality between social media engagement and career outcomes remains difficult to conclusively establish due to endogeneity issues. Further empirical validation through surveys can increase internal validity. Evolving platform algorithms that determine content visibility and popularity complicate historical data reliability. Continued monitoring is necessary to track the impacts of these algorithm shifts.

Future studies can address these limitations through cross-platform comparative analyses, investigating platform-specific and industry-specific success metric variations, establishing greater empirical causality between engagement and career advancement, and sustained tracking of algorithm changes. This can provide more generalized and culturally-nuanced principles for effectively navigating professional social media spaces.

Author contributions: PNU: conceptualization, data curation, formal analysis, methodology, writing – original draft; **NNU:** investigation, writing – original draft, writing – review & editing; **EVG:** formal analysis, investigation, writing – review & editing; **RLB:** Conceptualization, formal analysis, investigation, writing – original draft; **AVK:** data curation, formal analysis, methodology; **NNK:** investigation, methodology, writing – review & editing. All authors approved the final version of the article.

Funding: The authors received no financial support for the research and/or authorship of this article.

Ethics declaration: The authors declared that the study does not require approval from an ethics committee since data is retrieved from the social media application via APIs. There are no variables that would reveal the identities of the participants.

Declaration of interest: The authors declare no competing interest.

Data availability: Data generated or analyzed during this study are available from the authors on request.

REFERENCES

Amadoru, M., & Gamage, C. (2016). Evaluating effective use of social networks for recruitment. In *Proceedings of the 2016 ACM SIGMIS Conference on Computers and People Research* (pp. 125-133). ACM. https://doi.org/10.1145/2890602.2890604

Asad, R., Altaf, S., Ahmad, S., Shah Noor Mohamed, A., Huda, S., & Iqbal, S. (2023). Achieving personalized precision education using the Catboost model during the COVID-19 lockdown period in Pakistan. *Sustainability, 15*(3), 2714. https://doi.org/10.3390/su15032714

Baima, G., Forliano, C., Santoro, G., & Vrontis, D. (2020). Intellectual capital and business model: A systematic literature review to explore their linkages. *Journal of Intellectual Capital, 22*(3), 653-679. https://doi.org/10.1108/JIC-02-2020-0055

Bao, P. (2016). Modeling and predicting popularity dynamics via an influence-based self-excited Hawkes process. In *Proceedings of the International Conference on Information and Knowledge Management* (pp. 1897-1900). https://doi.org/10.1145/2983323.2983868

Blaer, M., Frost, W., & Laing, J. (2020). The future of travel writing: Interactivity, personal branding and power. *Tourism Management, 77*, 104009. https://doi.org/10.1016/j.tourman.2019.104009

Boyd, D. M., & Ellison, N. B. (2007). Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication, 13*(1), 210-230. https://doi.org/10.1111/j.1083-6101.2007.00393.x

Brandtzaeg, P. B., & Chaparro-Domínguez, M. Á. (2020). From youthful experimentation to professional identity: Understanding identity transitions in social media. *Young, 28*(2), 157-174. https://doi.org/10.1177/1103308819834386

Brodovskaya, E. V., Dombrovskaya, A. Y., Pyrma, R. V., Sinyakov, A. V., & Azarov, A. A. (2019). The impact of digital communication on Russian youth professional culture: Results of a comprehensive applied study. *Monitoring of Public Opinion: Economic and Social Changes, 1*, 228-251. https://doi.org/10.14515/monitoring.2019.1.11

Brown, C., Hooley, T., & Wond, T. (2020). Building career capital: Developing business leaders' career mobility. *Career Development International, 25*(5), 445-459. https://doi.org/10.1108/CDI-07-2019-0186

Buettner, R. (2017). Getting a job via career-oriented social networking markets: The weakness of too many ties. *Electronic Markets, 27*(4), 371-385. https://doi.org/10.1007/s12525-017-0248-3

Cakit, E., & Dagdeviren, M. (2022). Predicting the percentage of student placement: A comparative study of machine learning algorithms. *Education and Information Technologies, 27*(1), 997-1022. https://doi.org/10.1007/s10639-021-10655-4

Carpentier, M., Van Hoye, G., & Weijters, B. (2019). Attracting applicants through the organization's social media page: Signaling employer brand personality. *Journal of Vocational Behavior, 115*, 103326. https://doi.org/10.1016/j.jvb.2019.103326

Cerruto, F., Cirillo, S., Desiato, D., Gambardella, S. M., & Polese, G. (2022). Social network data analysis to highlight privacy threats in sharing data. *Journal of Big Data, 9*, 19. https://doi.org/10.1186/s40537-022-00566-7

Chen, X., Wei, S., Davison, R. M., & Rice, R. E. (2020). How do enterprise social media affordances affect social network ties and job performance? *Information Technology and People, 33*(1), 361-388. https://doi.org/10.1108/ITP-11-2017-0408

Donelan, H. (2016). Social media for professional development and networking opportunities in academia. *Journal of Further and Higher Education, 40*(5), 706-729. https://doi.org/10.1080/0309877X.2015.1014321

Donskoy, A. G., Sakhno, O. A., & Makashova, V. N. (2021). Professional network communities as a resource for informal professional development of teaching staff. *Scientific Support of the Personnel Development System, 2*(47), 15-30.

Duffy, B. E., & Pooley, J. D. (2017). “Facebook for academics”: The convergence of self-branding and social media logic on academia.edu. *Social Media and Society, 3*(1). https://doi.org/10.1177/2056305117696523

El Ouirdi, M., Segers, J., El Ouirdi, A., & Pais, I. (2015). Predictors of job seekers’ self-disclosure on social media. *Computers in Human Behavior, 53*, 1-12. https://doi.org/10.1016/j.chb.2015.06.039

Enughwure, A. A., & Ogbise, M. E. (2020). Application of machine learning methods to predict student performance: A systematic literature review. *International Research Journal of Engineering and Technology, 7*(5), 3405-3415.

Jacobson, J. (2020). You are a brand: Social media managers’ personal branding and “the future audience.” *Journal of Product and Brand Management, 29*(6), 715-727. https://doi.org/10.1108/JPBM-03-2019-2299

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of social media. *Business Horizons, 53*(1), 59-68. https://doi.org/10.1016/j.bushor.2009.09.003

Khamis, S., Ang, L., & Welling, R. (2017). Self-branding, ‘micro-celebrity’ and the rise of social media influencers. *Celebrity Studies, 8*(2), 191-208. https://doi.org/10.1080/19392397.2016.1218292

Koch, T., Gerber, C., & De Klerk, J. J. (2018). The impact of social media on recruitment: Are you LinkedIn? *SA Journal of Human Resource Management, 16*, a861. https://doi.org/10.4102/sajhrm.v16i0.861

Koltsova, O., Koltcov, S., & Sinyavskaya, Y. (2017). When internet really connects across space: Communities of software developers in Vkontakte social networking site. In G. Ciampaglia, A. Mashhadi, & T. Yasseri (Eds.), *Social informatics* (pp. 431-442). https://doi.org/10.1007/978-3-319-67256-4_34

Kostin, D. V., & Shelukhin, O. I. (2016). Comparative analysis of machine learning algorithms for classification of encrypted network traffic. *T-Comm-Telecommunications and Transport, 10*(9), 43-52.

Kucharska, W. (2021). Leadership, culture, intellectual capital and knowledge processes for organizational innovativeness across industries: The case of Poland. *Journal of Intellectual Capital, 22*(7), 121-141. https://doi.org/10.1108/JIC-02-2021-0047

Kuiper, K. (2021). Communication theory of identity: A fifth frame. *Annals of the International Communication Association, 45*(3), 175-187. https://doi.org/10.1080/23808985.2021.1976069

Labrecque, L. I., Markos, E., & Milne, G. R. (2011). Online personal branding: Processes, challenges, and implications. *Journal of Interactive Marketing, 25*(1), 37-50. https://doi.org/10.1016/j.intmar.2010.09.002

Lardo, A., Dumay, J., Trequattrini, R., & Russo, G. (2017). Social media networks as drivers for intellectual capital disclosure: Evidence from professional football clubs. *Journal of Intellectual Capital, 18*(1), 63-80. https://doi.org/10.1108/JIC-09-2016-0093

Leite, F. P., & de Baptista, P. P. (2022). The effects of social media influencers’ self-disclosure on behavioral intentions: The role of source credibility, parasocial relationships, and brand trust. *Journal of Marketing Theory and Practice, 30*(3), 295-311. https://doi.org/10.1080/10696679.2021.1935275

Lin, J., Luo, Z., Cheng, X., & Li, L. (2019). Understanding the interplay of social commerce affordances and swift guanxi: An empirical study. *Information and Management, 56*(2), 213-224. https://doi.org/10.1016/j.im.2018.05.009

Luwie, L., & Pasaribu, L. H. (2021). The influence of personal branding in the establishment of social media influencer credibility and the effect on brand awareness and purchase intention. *Enrichment: Journal of Management, 12*(1), 917-925.

Macy, M. W. (2023). The antecedents and consequences of network mobility. *PNAS, 120*(28), e2306897120. https://doi.org/10.1073/pnas.2306897120

Mahmutova, D., & Gerasimova, D. (2023). Digital illustration marketing via blogs in the VKontakte social network. *Virtual Communication and Social Networks, 2023*(3), 160-166. https://doi.org/10.21603/2782-4799-2023-2-3-160-166

Marwick, A. E., & Boyd, D. (2011). I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New Media and Society, 13*(1), 114-133. https://doi.org/10.1177/1461444810365313

Merunková, L., & Šlerka, J. (2019). Goffman´s theory as a framework for analysis of self-presentation on online social networks. *Masaryk University Journal of Law and Technology, 13*(2), 243-276. https://doi.org/10.5817/MUJLT2019-2-5

Mogaji, E. (2019). Student engagement with LinkedIn to enhance employability. In A. Diver(Ed.), *Employability via higher education: Sustainability as scholarship* (pp. 321-329). Springer. https://doi.org/10.1007/978-3-030-26342-3_21

Nikitkov, A., & Sainty, B. (2014). The role of social media in influencing career success. *International Journal of Accounting and Information Management, 22*(4), 273-294. https://doi.org/10.1108/IJAIM-02-2014-0009

Orishev, A. B., Mamedov, A. A., Kotusov, D. V., Grigoriev, S. L., & Makarova, E. V. (2020). Digital education: VKontakte social network as a means of organizing the educational process. *Journal of Physics: Conference Series, 1691*, 012092. https://doi.org/10.1088/1742-6596/1691/1/012092

Osisanwo, F. Y., Akinsola, J. E. T., Awodele, O., Hinmikaiye, J. O., Olakanmi, O., & Akinjobi, J. (2017). Supervised machine learning algorithms: Classification and comparison. *International Journal of Computer Trends and Technology, 48*(3), 128-138. https://doi.org/10.14445/22312803/IJCTT-V48P126

Pennycook, G., Epstein, Z., Mosleh, M., Arechar, A. A., Eckles, D., & Rand, D. G. (2021). Shifting attention to accuracy can reduce misinformation online. *Nature, 592*(7855), 590-595. https://doi.org/10.1038/s41586-021-03344-2

Popov, E. V., Simonova, V. L., & Komarova, O. V. (2021). The regional and substantial differentiation of social media. *Applied Economics Letters, 28*(16), 1386-1390. https://doi.org/10.1080/13504851.2020.1820435

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018). CatBoost: Unbiased boosting with categorical features. In *Proceedings of the 32nd Conference on Neural Information Processing Systems* (pp. 1-11).

Putnam, R. D. (2000). *Bowling alone: The collapse and revival of American community*. Simon and Dchuster. https://doi.org/10.1145/358916.361990

Ramaswami, G., Susnjak, T., & Mathrani, A. (2022). On developing generic models for predicting student outcomes in educational data mining. *Big Data and Cognitive Computing, 6*(1), 6. https://doi.org/10.3390/bdcc6010006

Ruparel, N., Dhir, A., Tandon, A., Kaur, P., & Islam, J. U. (2020). The influence of online professional social media in human resource management: A systematic literature review. *Technology in Society, 63*, 101335. https://doi.org/10.1016/j.techsoc.2020.101335

Sheer, V. C., & Rice, R. E. (2017). Mobile instant messaging use and social capital: Direct and indirect associations with employee outcomes. *Information and Management, 54*(1), 90-102. https://doi.org/10.1016/j.im.2016.04.001

Shetty, S. H., Shetty, S., Singh, C., & Rao, A. (2022). Supervised machine learning: Algorithms and applications. In P. Singh (Ed.), *Fundamentals and methods of machine and deep learning* (pp. 1-16). Wiley. https://doi.org/10.1002/9781119821908.ch1

Singh, A. P. (2023). A study on impact of social media on recruitment process. *International Journal of Scientific Research in Engineering and Management*. https://doi.org/10.55041/IJSREM18787

Tomasevic, N., Gvozdenovic, N., & Vranes, S. (2020). An overview and comparison of supervised data mining techniques for student exam performance prediction. *Computers and Education, 143*, 103676. https://doi.org/10.1016/j.compedu.2019.103676

Tsaturyan, M. M., & Matevosyan, A. P. (2022). Features of the functioning of the virtual communication language in the social network "VKontakte." *Bulletin of the Adyghe State University, Series "Philology and Art History," 2*(297), 107-115. https://doi.org/10.53598/2410-3489-2022-2-297-107-115

Umair, A., Masciari, E., & Ullah, M. H. (2023). Vaccine sentiment analysis using BERT + NBSVM and geo-spatial approaches. *Journal of Supercomputing, 79*(15), 17355-17385. https://doi.org/10.1007/s11227-023-05319-8

Van Den Beemt, A., Thurlings, M., & Willems, M. (2020). Towards an understanding of social media use in the classroom: A literature review. *Technology, Pedagogy and Education, 29*(1), 35-55. https://doi.org/10.1080/1475939X.2019.1695657

van Dijck, J. (2013). "You have one identity": Performing the self on Facebook and LinkedIn. *Media, Culture and Society, 35*(2), 199-215. https://doi.org/10.1177/0163443712468605

Van Osch, W., & Steinfield, C. W. (2018). Strategic visibility in enterprise social media: Implications for network formation and boundary spanning. *Journal of Management Information Systems, 35*(2), 647-682. https://doi.org/10.1080/07421222.2018.1451961

Van Zoonen, W., Verhoeven, J. W. M., & Vliegenthart, R. (2016). How employees use Twitter to talk about work: A typology of work-related tweets. *Computers in Human Behavior, 55*, 329-339. https://doi.org/10.1016/j.chb.2015.09.021

Vartanov, S., & Khvorostyanaya, A. (2023). Personal brand strategizing in digital mediatization: Game-theoretic and behavioral approaches. *Strategizing: Theory and Practice, 3*(2), 218-233. https://doi.org/10.21603/2782-2435-2023-3-2-218-233

Wang, X., Zhang, R., Wang, X., Xu, D., & Tian, F. (2022). How do mobile social apps matter for college students' satisfaction in group-based learning? The mediation of collaborative learning. *Frontiers in Psychology, 13*. https://doi.org/10.3389/fpsyg.2022.795660

Xie, X., & Liu, L. (2022). Exploring the antecedents of trust in electronic word-of-mouth platform: The perspective on gratification and positive emotion. *Frontiers in Psychology, 13*. https://doi.org/10.3389/fpsyg.2022.953232

Xu, Z., & Qian, M. (2023). Predicting popularity of viral content in social media through a temporal-spatial cascade convolutional learning framework. *Mathematics, 11*(14), 3059. https://doi.org/10.3390/math11143059