www.igi-global.com

Statistical Hypothesization and Predictive Modeling of Reactions to COVID-19 Induced Remote Work

A study to understand the general trends of response to pursuing academic and professional commitments virtually.

Arjun Sharma
School of Electronics Engineering, Vellore Institute of Technology (VIT), Chennai, India – 600127

Hemanth Harikrishnan
School of Electronics Engineering, Vellore Institute of Technology (VIT), Chennai, India - 600127

Sathiya Narayanan
School of Electronics Engineering, Vellore Institute of Technology (VIT), Chennai, India - 600127

Om Prakash Swain
School of Electronics Engineering, Vellore Institute of Technology (VIT), Chennai, India - 600127

Utkarsh
School of Electronics Engineering, Vellore Institute of Technology (VIT), Chennai, India - 600127

Akshay Giridhar
School of Electronics Engineering, Vellore Institute of Technology (VIT), Chennai, India - 600127

ABSTRACT

The initial outbreak of the Coronavirus was met with lockdowns being enforced all over the world in March, 2020. A prominent change in human lifestyle is the shift of professional and academic work to online platforms, as opposed to previously attending to them in person. As with any major change, the implementation of complete remote work and study is expected to affect different people differently. Through the results of a questionnaire designed as per the implications of the self-efficacy theory shared with people who were either students, working professionals, entrepreneurs, and homemakers aged between 12 and 60 years, the authors perform statistical analysis and subsequently hypothesize how different aspects of remote work affect the population from a mental standpoint using t-test, with respect to their professional or academic work. This is followed by predictive modelling through machine learning algorithms to classify working preference for as 'remote' or 'in-person'.

Keywords: COVID-19, t-test, Supervised Learning, Self-Efficacy, Machine Learning, Data Analysis, Correlation.

INTRODUCTION

COVID-19 has spread across the world, with the World Health Organization terming it a 'pandemic' in March, 2020 (World Health Organization, 2020). It has infected over 238 million across the globe, killing at least 4.8 million as of October, 2021 (Worldometers.info. 2021). To reduce the spread and impact of the disease, government sanctioned lockdowns were imposed. The lockdowns were accompanied by the adoption of remote work and study alternatives.
This work focuses on understanding the effect of various aspects of complete remote work through means of a survey answered by 450 respondents aged between 12 and 60 years across India. At the time of data collection, all the participants were WFH full-time. Participation in the research was voluntary, anonymous, and without any reward. The process of data collection was in full compliance of the declaration of Helsinki (World Medical Association Declaration of Helsinki, 2013). The questionnaire comprises of questions modelled on the basis of the implications of the self-efficacy theory and its 4 main constructs described in section II. The data obtained is analyzed to understand factors that influence an individual's ability to work, how their state of mind was affected by new norms, and also their general experience in a virtual setting. Using t-test, the proposed hypotheses are validated for different behavioral aspects showcased by respondents. Finally, predictive modelling using 6 supervised learning models is performed to predict an individual's preferred method of working as either 'remote', or 'in-person'. In this work, section II discusses the Self-Efficacy Theory and other relevant research undertakings. Section III discusses the methods involved for questionnaire preparation. Section IV comprises the exploratory data analysis of the survey results. Section V constitutes the hypotheses proposed by the authors regarding the behavioral experiences of those working remote. Section VI summarizes the results obtained in predictive modelling stage. Section VII concludes the work.
The contributions through this work are, firstly, formulating and testing hypotheses the results of which, can be used to potentially increase the quality of work and level of satisfaction of employees working remotely. Secondly, the usage of predictive analytics in determining the work-pattern (remote or non-remote) preferences of individuals from the set of features discussed in section V.

BACKGROUND

Self-efficacy theory (Bandura, 1986) claims that the behavior, environment, and cognitive factors of an individual share high levels of interrelation. Self-efficacy is defined as “a judgement of one’s ability to execute a particular behavior pattern.” Further assessment shows that self-efficacy is crucial in levels of motivation and performance exhibited by an individual (Wood & Bandura, 1989), implying that self-efficacy levels also influence the amount of effort and/or time invested on a particular task. Those with higher beliefs of self-efficacy take more effort to complete their tasks, conversely, those with lower self-efficacy beliefs tend to undertake relatively lesser effort, spend less time and sometimes even abandoning it.

According to self-efficacy theory, 4 major sources are considered by an individual through the formation of their self-efficacy judgements, they are: Performance Accomplishment, Vicarious Experience, Social Persuasion and, Physiological and Emotional State (Wrycza & Maslankowski, 2020).

Self-efficacy theory excels in behavior and performance prediction (Bandura, 1978). The theory and evidence supporting its empirical implications are robust and its implications strongly suit the study of virtual organizations (Staples et al., 1999). Employees working remote have little support or guidance, strongly fitting the current situation across the world due to COVID-19.

In addition, the authors have compared their work with that of other research groups. In (Zhang C et al., 2021; Galanti T et al., 2021), the general opinion of the public regarding remote-work due to COVID-19 enforced lockdowns was gathered through the online microblogging website “Twitter.” Over 500,000 ‘tweets’ (Microblogs of less than 140 characters) were analyzed. They perform time-series analysis to plot the frequency of remote-work related tweets over time. Sentiment analysis is performed to understand to what extent remote-work is embraced. Over 50% of the tweets are in favor of remote work, about 40% tweets have a neutral stance, and about 7.5% tweets are negative. While the study highlights the general proportion of sentiment in favor or against remote work, it does not highlight such intricacies across other demographics such as age or gender. It also does not account for different work patterns and influence of remote-work on the mindset of people.

The authors in (World Health Organization, 2021) model the intricacies of daily life while working from home by exploring the influence of family-work conflict, social isolation, distracting environment, job autonomy, and self-leadership have on employees’ productivity, work engagement, and stress. However, this study does not include anyone who is not an employee such as students, entrepreneurs, etc. In addition, the study also does not involve a means to predict one’s preference of working in a particular setting, over others.

Overall, the existing literature either does not highlight the experience of individual other than employees, in addition, they not involve a predictive modelling approach towards determining factors that could or could not contribute to the preferences of working from home or working in-person which are highlighted in this work.

METHOD FOR QUESTIONNAIRE FRAMING

The appendix of this work has the questions that were part of the questionnaire that was shared with participants. While the self-efficacy theory broadly covers possible questions that participants may be asked, it was important to design questions based on existing evidence for potential of altering one’s approach to working remote. Age, gender, and occupation (i, ii, iii in appendix respectively), were asked to establish trends across demographics during the exploratory data analysis phase. In addition, the considerations given by WHO (Chung et al., 2020), were also taken into consideration for question formation.

Research undertaken by other research groups shows that professionals tend to spend more hours working while working from home (Gibbs et al., 2021), to verify the occurrence of the same, questions iv and v (in appendix) were asked. Time spent travelling was considered part of time spent working for the reason that this travel on a daily basis was to pursue their professional commitments in their office/institution.

In regards to productivity levels, other research groups (World Health Organization, 2020; Zhang C et al, 2021) report a decrease in general productivity among the working population. Productivity being an important metric for measuring self-efficacy, makes question vii important in WFH context.

As specified by (Chung et al., 2020; Gibbs et al, 2021), sleep cycles and patterns were generally observed to be disturbed. Question viii was imperative to understand the sleep patterns changes (if any), in the general opinion of the answerers of the questionnaire.

With the lockdown, the concerns on mental health, picking up hobbies, in addition to family interactions, and taking care of one's own well-being would be important considerations (Yuksel D et al., 2021; Conroy DA et al., 2021; Morse KF et al., 2021). These evidences make it necessary to ask questions ix, x, xi, xii and xvi (in appendix).

The complete migration to virtual platforms requires getting acclimatized to the online platforms and media of communication online. Considering this and those in (Morelli M et al., 2021), it necessitates asking the question xiii (in appendix).

Another important consideration is whether the home environment allows one to focus on their commitments virtually, (Rehman et al., 2021) mentions the possibility of distractions at home not allowing one to function with appreciable focus, making question xiv important from WFH standpoint.

To understand the overall preference in work setting (either remote or in-person), question xxii (in appendix), was asked. This would also serve as the target variable for the predictive analytics procedure as described in section VI.

EXPLORATORY DATA ANALYSIS

1. Daily Time Commitment

Table I represents age wise trends on the per-day average time spent on work. The age group between 12 to 18 years old shows a decrease of 1.7 hours of time commitment per day. This can be attributed to zero travel time and less time preparing for school. A similar trend is observed for those between ages 19 and 25, as seen in Table 1. Another reason for a sharp decrease in time spent working on average could be the result of widespread loss of employment, especially for those between ages 18 and 24 (Gould et al., 2020).

Respondents between 26 and 32 years on average, spent more time working at home than they spent travelling and working combined on a regular day. It is observed that except for students, everyone else spent more time working at home. Notably, homemakers saw the largest increase in their working hours during the pandemic.

Table 1. Daily time commitment across ages before and during the pandemic

Age Group	Pre-Pandemic Work Hours (Work + Travel)	Work Hours at Home
12-18	7.4	5.7
19-25	7.9	6.8
26-32	8.7	9.4
33-40	8.3	7.4
41-50	9.6	8.7
51-60	8	9.45

2. **Adjusting to New Setting**

Respondents rated their ease in adjusting to an online environment from 1 to 5 in increasing order of difficulty. From Table 2, males found it tougher migrating to online environment than females except for those in aged between 19-25 and 26-32. Those between 26-32 years found it easiest to migrate to online platforms for both genders. Said age group corresponds to a relatively early stage of one's career, implying that those in early career stages are more adept at working online, while those aged between 33-40 years found it most difficult to make the same adjustment for both genders. From Table 3, homemakers found it the easiest to adapt to an online work environment, followed by entrepreneurs. Students in general experience difficulty.

Table 2. Trends across gender and age pertaining to the ease in adjusting to an online setting

Age	Gender	Ease of Adjustment
12-18	Male	3.41
	Female	2.64
19-25	Male	2.48
	Female	2.79
26-32	Male	1.91
	Female	1.97
33-40	Male	3.56
	Female	3.94
41-50	Male	2.32
	Female	1.87
51-60	Male	2.80
	Female	2.66

Table 3. Trends across occupation pertaining to the ease in adjusting to an online setting

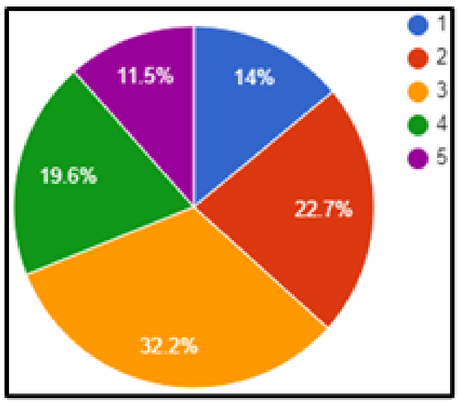
Occupation	Ease of Adjustment

Student in School	2.61
Student in College	2.73
Entrepreneur	2.08
Homemaker	1.49
Working Professional	2.43

3. Effect of Home Environment

Respondents rated the level of disturbance they encountered during work hours from 1 to 5 in increasing order of noisiness, where 1 signified a quiet environment, and 5 signified a highly chaotic and noisy environment. From Figure 1, the level of disturbance experienced by respondents follows a normal distribution with mean 2.91 (implying a mostly neutral outlook towards level of noise at home), and standard deviation 1.198.

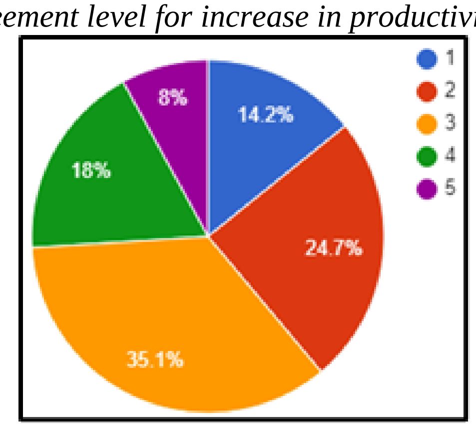
Figure 1. Disturbance level distribution experienced by respondents working remote



4. Changes in Productivity

Respondents rated their level of productivity while working remote from 1 to 5 In increasing order. 1 signified no productivity change, and 5 signified substantial increase. From Figure 2, there is a slight disagreement to this statement. The responses follow a normal distribution with mean 2.81, and standard deviation of 1.132.

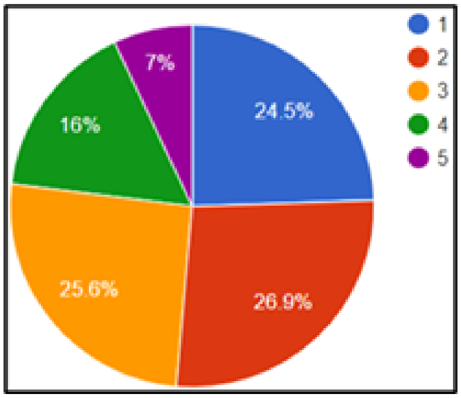
Figure 2. Distribution of agreement level for increase in productivity



5. **Changes in Sleep Cycle**

Respondents rated their level of agreement to whether their sleep cycle has improved working remote from 1 to 5 in increasing order of agreement. As per Figure 3, there is disagreement. The responses follow a normal distribution with mean 2.54 and standard deviation 1.218.

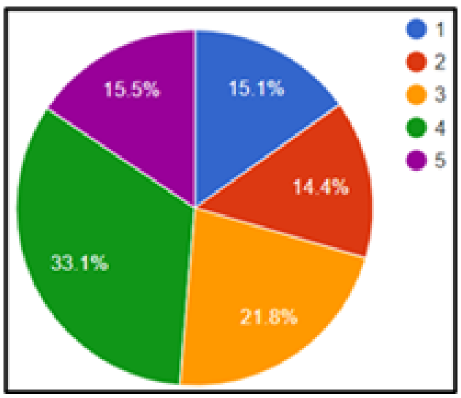
Figure 3. Distribution of agreement level for change in sleep cycle



6. **New Hobbies for Recreation**

Respondents rated their level of agreement with whether they were able to pick up new hobbies while working remote from 1 to 5 in increasing order of agreement. As per Figure 4, there is agreement to this statement. The response follows a normal distribution, with a mean of 3.2, and standard a deviation of 1.289.

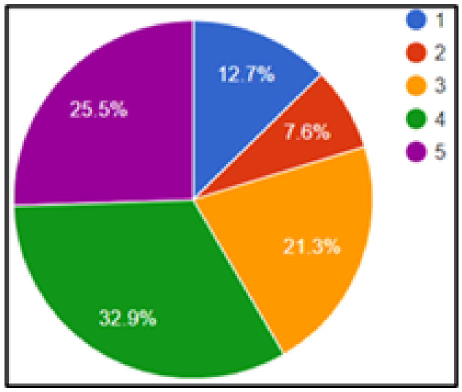
Figure 4. Distribution of agreement level for picking up new hobbies



7. **Time Spent with Family**

Respondents rated their level of agreement with whether they were spending more time with family while working remote from 1 to 5 in increasing order of agreement. From Figure 5, there is strong agreement. The response follows a normal distribution, with mean 3.51, and a standard deviation of 1.293.

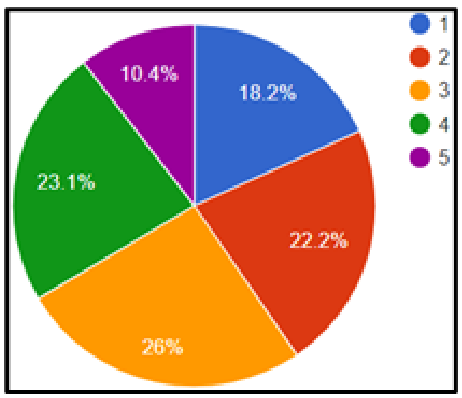
Figure 5. Distribution of agreement level for time spent with family



8. Stress Levels

Respondents rated their level of agreement with whether they experience lower stress while working remote from 1 to 5 in increasing order of agreement. As per Figure 6, the response indicates slight disagreement, following a normal distribution, with a mean of 2.85, and a standard deviation of 1.256.

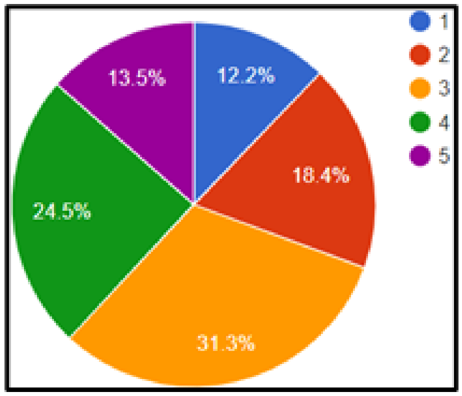
Figure 6. Distribution of agreement level for Stress level reduction



9. Time Spent on Self-Care

Respondents rated their level of agreement with whether they were spending time on their self-care while working remote from 1 to 5 in increasing order of agreement. As per Figure 7, there is agreement. The response follows a normal distribution, with a mean of 3.22, and a standard deviation of 1.445.

Figure 7. Distribution of agreement level for spending time on self-care



10. Overall Working Pattern Preference

Respondents were asked to choose their overall working pattern preference from either working/studying remotely or working/studying in-person (physically attending to their work or their classes). From Figure 8 it is evident that the majority of the respondents prefer to go to their workplace or institution and attend to their commitments in-person, rather than pursuing it online, with 72.2% respondents preferring in-person work/classes, and 27.8% preferring to do so virtually.

Figure 8. Distribution of overall working preference (remote or in-person)

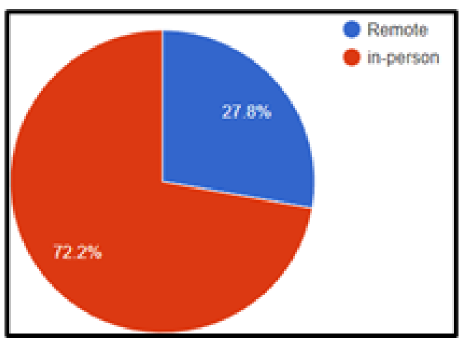


Table 4 highlights correlation observed between constructs that influence self-efficacy in a remote-work setting from the data obtained from the answers given by respondents to the questionnaire. Table 4 presents 8 metrics and their mutual level of correlation. Based on the correlation levels for any 2 metrics, those with significant absolute values are chosen. 2 tailed paired t-test is performed for validation of Hypothesis at p=0.05.

Table 4. Table representing correlation between different behavioral constructs while pursuing occupational commitments virtually

	1.	2.	3.	4.	5.	6.	7.	8.
1. Difficulty in Adjusting to Online Work	1	0.51	-0.64	-0.054	-0.24	-0.14	-0.26	-0.21
2. Noise in Home Environment	0.51	1	-0.37	-0.24	-0.18	-0.18	-0.26	-0.19
3. Productivity Working Remote	-0.64	-0.37	1	0.14	0.31	0.21	0.29	0.23

4. Balanced Sleep Hours	-0.054	-0.24	0.14	1	0.56	0.56	0.63	0.55
5. Picking up Hobbies	-0.24	-0.18	0.31	0.56	1	0.71	0.69	0.69
6. Time Spent with Family	-0.14	-0.18	0.21	0.56	0.71	1	0.68	0.71
7. Tendency to lower stress	-0.26	-0.26	0.29	0.63	0.69	0.68	1	0.72
8. Time spent on self-care	-0.21	-0.19	0.23	0.55	0.69	0.71	0.72	1

Hypothesis One: Increased Difficulty in Adjusting to Online Environment is associated with Noisy Home Environments

From Table 4, the difficulty in adjusting to an online environment seems to share a linear relationship with the level of noise at home, with a correlation coefficient of **0.51**. Consider:

H_0: There is no significant evidence for a linear relationship between increased difficulty in adjusting to an online environment and noisy home environment.

H_A: There is significant evidence for a linear relationship between increased difficulty in adjusting to an online environment and noisy home environment.

Hypothesis test outcome:

t	df	p	95% CI
1.4747	449	0.1412	-0.03 to – 0.22

Since $p>0.05$, H_A is rejected and H_0 is Accepted. There is no significant evidence for a linear relationship between noisy home environment and difficulty in adjusting to an online environment.

Hypothesis Two: Picking up hobbies is associated with balanced sleep when working remote

From Table 4, building new hobbies seems to share a linear relationship with balanced sleep, with a correlation coefficient of **0.56**. Consider:

H_0: There is no significant evidence for a linear relationship between picking up hobbies and balanced sleep.

H_A: There is significant evidence for a linear relationship between picking up hobbies and balanced sleep.

Hypothesis test outcome:

t	df	p	95% CI
9.3672	449	$3.6354e^{-19}$	0.51 to 0.79

Since $p<0.05$, H_0 is rejected and H_A is Accepted. There is significant evidence for a linear relationship between picking up hobbies and balanced sleep.

Hypothesis Three: Time Spent with family is associated with picking up hobbies when working remote

From Table 4, time spent with family seems to share a linear relationship with picking up hobbies, with a correlation coefficient of **0.71**. Consider:

H_0: There is no significant evidence for a linear relationship between time spent with family and picking up hobbies.

H_A: There is significant evidence for a linear relationship between time spent with family and picking up hobbies.

Hypothesis test outcome:

t	df	p	95% CI
5.7946	449	$1.2907e^{-8}$	0.21 to 0.42

Since p<0.05, H0 is rejected and HA is Accepted. There is significant evidence for a linear relationship between time spent with family and picking up hobbies.

Hypothesis Four: Time spent with family is associated with Balanced Sleep when working remote

From Table 4, time spent with family seems to share a linear relationship with balanced sleep cycles, with a correlation coefficient of **0.56**. Consider:

H_0: There is no significant evidence for a linear relationship between time spent with family and having a balanced sleep.
H_A: There is significant evidence for a linear relationship between time spent with family and having a balanced sleep.

Hypothesis test outcome:

t	df	p	95% CI
9.5543	449	$1.7868e^{-39}$	0.672 to 1.012

Since p<0.05, H0 is rejected and HA is Accepted. There is significant evidence for a linear relationship between time spent with family and having a balanced sleep.

Hypothesis Five: Time spent on self-care is associated with lower stress when working remote

From Table 4, time spent on self-care seems to share a linear relationship with lower stress levels, with a correlation coefficient of **0.72**. Consider:

H_0: There is no significant evidence for a linear relationship between time spent on self-care and lower stress levels.
H_A: There is significant evidence for a linear relationship between time spent on self-care and lower stress levels.

Hypothesis test outcome:

t	df	p	95% CI

5.0201	449	$7.4597e^{-7}$	0.14 to 0.32

Since $p<0.05$, H0 is rejected and HA is Accepted. There is significant evidence for a linear relationship between time spent on self-care and lower stress levels.

Hypothesis Six: Low Levels of Stress are related with a Balanced Sleep when working remote

From Table 4, lower stress seems to share a linear relationship with a balanced sleep, with a correlation coefficient of **0.63**. Consider:

H_0: There is no significant evidence for a linear relationship between lower stress levels and balanced sleep.
H_A: There is significant evidence for a linear relationship between lower stress levels and balanced sleep.

Hypothesis test outcome:

t	df	p	95% CI
4.9272	449	$2.377e^{-16}$	0.19 to 0.43

Since $p<0.05$, H0 is rejected and HA is Accepted. There is significant evidence for a linear relationship between lower stress levels and balanced sleep.

Hypothesis Seven: Picking up Hobbies is Associated with Lower Stress when working remote

From Table 4, picking up hobbies while working remote sems to share a linear relationship with lower stress levels, with a correlation coefficient of **0.69**. Consider:

H_0: There is no significant evidence for a linear relationship between picking up hobbies and lower stress levels.
H_A: There is significant evidence for a linear relationship between picking up hobbies and lower stress levels.

Hypothesis test outcome:

t	df	p	95% CI
3.5612	449	$1.139e^{-8}$	0.20 to 0.69

Since $p<0.05$, H0 is rejected and HA is Accepted. There is significant evidence for a linear relationship between picking up hobbies and lower stress levels.

Hypothesis Eight: Time spent on self-care is associated with picking up Hobbies when working remote

Table 4 suggests that time spent on self-care shares a linear relationship with picking up new hobbies, with a correlation coefficient of **0.69**. Consider:

H_0: There is no significant evidence for a linear relationship between time spent on self-care and picking up hobbies.

H$_A$: There is significant evidence for a linear relationship between time spent on self-care and picking up hobbies.

Hypothesis test outcome:

t	df	p	95% CI
2.2442	449	0.0253	0.01to 0.20

Since $p<0.05$, H0 is rejected and HA is Accepted. There is significant evidence for a linear relationship between time spent on self-care and picking up hobbies.

MACHINE LEARNING BASED PREDICTIVE MODELLING

1. Data Preprocessing

Data obtained from respondents is encoded through the Scikit-learn library in Python (Pedregosa et al., 2011). The categorical features were One-Hot encoded; different categorical values were converted to numerical values, further normalized to the range from -1 to 1, minimizing influence of outliers in the predictions. With respect to dependent features Label encoder was used as the values are cardinal. To train the model, the data was split into train and test set. To reduce the effect outliers Standard scaling is performed. Next, Principal component analysis is used to reduce the dimensions of independent features. The dataset does not contain duplicate or missing values because all questions were compulsory to answer, and respondents could answer the survey only once.

2. Work Setting Preference Model

The prediction of preference of remote work or working in-person is a binary classification problem, and hence requires the use of supervised learning models. The interest in the implementation of these models is to detect as to whether or not the individual would prefer working/studying remote over studying/working in-person. The following classifier models were used: Logistic Regression, K Nearest Neighbor, Random Forest, Naïve Bayes, XGBoost, Passive Aggressive Classifier, and Support Vector Classifier, also using the scikit-learn library. The dataset was split into a ratio of 80:20 corresponding to training and testing sets respectively.

3. Machine Learning Based Predictive Models

i. **Logistic Regression**: A linear model for binary classification using the logistic function to model the dependent variables (King & Zeng, 2001).

ii. **Random Forest:** A tree-based bagging algorithm where the successive trees are created by usage of different samples in the dataset. For prediction purposes, the simple majority vote is taken (Liaw & Wiener, 2002).

iii. **Naïve Bayes:** It is a supervised machine learning model used for classification tasks. It is a probabilistic model based on Bayes Theorem (Webb G.I., 2011).

iv. **XGBoost:** A tree-based gradient boosting algorithm uses a set of weak learners. After fitting these weak learners, the final prediction is produced by the combination of the predictions of the weak learners through weighted sum (Chen & Guestrin, 2016).

v. **Passive Aggressive Classifier:** This is an online learning model, making it very effective in situations with continuous data input. It remains passive in case of a correct prediction, and responds aggressively to incorrect predictions (Crammer et al., 2006).

vi. **Support Vector Classifier**: An algorithm used for classification that may or may not be linear. The algorithm uses hyperplanes in higher dimensional space to perform class separation (Smola & Schölkopf, 2004).

4. Experimentation Results

The implemented models are compared to ascertain the most appropriate for classifying between virtual and in-person work preference. The Model uses 18 features: Age, Gender, Occupation, Daily Work Hours Before Pandemic, Daily Work Hours During Pandemic, Average Daily Travel Time, Difficulty in Adjusting to Online Environment, Home Noise Levels, Productivity at Home, Balanced Sleep Levels, New Hobbies, Time Spent with Family, Stress Level, Time Spent on Self-Care, Appealing Factors of Remote Work, Negatives of Remote Work, and overall work preference. As in Table VI, XGBoost Classifier model performs best for the dataset gathered through the survey in terms of accuracy and the F1 score.

Table 5. Performance evaluation of the models using the real dataset

Classifier Model	Accuracy (%)	F1 Score
Logistic Regression	92.34	0.92
Random Forest	83.83	0.86
Naïve Bayes	88.09	0.88
XGBoost	**94.04**	**0.94**
Passive Aggressive	91.49	0.92
Support Vector	93.19	0.93

CONCLUSION

This study ascertains different influences on one's mindset in a remote setting using statistical analysis techniques and predictive modelling. Contributions can be summarized in three points: 1) Proposal of hypotheses that relate different aspect of remote work and how they influence an individual's mindset, 2) Predictive modelling to ascertain the tendency of preferring one form of work over another, and 3) The usage of the results of the above hypotheses to alter to an extent, the approach in such spheres to create a productive atmosphere.

FUTURE RESEARCH DIRECTIONS

This study was aimed at the general population to summarize the response towards having to pursue academic and professional commitments from a remote setting. While the authors believe that this work is effective in its objective, there is tremendous scope to perform a similar study by focusing on specific groups of interests in this scenario such as medical professionals and other essential workers whose lifestyle is vastly different from others in the current scenario surrounding the pandemic.

ACKNOWLEDGMENT

The authors would like to express their deepest gratitude to the respondents of the questionnaire, without the data contributed by them, this study and its implications would not have been possible. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

REFERENCES

World Health Organisation (2020, March 11). Coronavirus disease 2019 (COVID-19) Situation Report – 51 https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200311-sitrep-51-covid-19.pdf?sfvrsn=1ba62e57_10.

Worldometers.info, 10 August, 2021, Dover, Delaware, USA

WMA Declaration of Helsinki, https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/

Bandura, A (1986), The Explanatory and Predictive Scope of Self-Efficacy Theory, Journal of Social and Clinical Psychology, Vol 4, No i l')Ht>, pp SS9-573.

Wood, Robert, and Albert Bandura. "Social Cognitive Theory of Organizational Management." The Academy of Management Review, vol. 14, no. 3, 1989, pp. 361–384. JSTOR, www.jstor.org/stable/258173.

Stanisław Wrycza & Jacek Maślankowski (2020) Social Media Users' Opinions on Remote Work during the COVID-19 Pandemic. Thematic and Sentiment Analysis, Information Systems Management, 37:4, 288-297, DOI: 10.1080/10580530.2020.1820631

Albert Bandura, Reflections on self-efficacy, Advances in Behaviour Research and Therapy, Volume 1, Issue 4, 1978, Pages 237-269, ISSN 0146-6402.

D. Sandy Staples, John S. Hulland, Christopher A. Higgins, (1999) A Self-Efficacy Theory Explanation for the Management of Remote Workers in Virtual Organizations. Organization Science 10(6):758-776. http://dx.doi.org/10.1287/orsc.10.6.758.

Zhang C, Yu MC, Marin S. Exploring public sentiment on enforced remote work during COVID-19. J Appl Psychol. 2021 Jun;106(6):797-810. doi: 10.1037/apl0000933. PMID: 34138587.

Galanti T, Guidetti G, Mazzei E, Zappalà S, Toscano F. Work From Home During the COVID-19 Outbreak: The Impact on Employees' Remote Work Productivity, Engagement, and Stress. J Occup Environ Med. 2021;63(7):e426-e432. doi:10.1097/JOM.0000000000002236

World Health Organization (2020a). Mental Health and Psychosocial Considerations During the COVID-19 Outbreak. Avaliable at: https://www.who.int/docs/default-source/coronaviruse/mental-health-considerations.pdf?sfvrsn=6d3578af_2 (March 18, 2020).

Working from home during the COVID-19 lockdown: Changing preferences and the future of work, : Heejung Chung, Hyojin Seo, Sarah Forbes, and Holly Birkett, DPC 130423 07/20, https://www.birmingham.ac.uk/Documents/college-social-sciences/business/research/wirc/epp-working-from-home-COVID-19-lockdown.pdf

Gibbs, Michael and Mengel, Friederike and Siemroth, Christoph, Work from Home & Productivity: Evidence from Personnel & Analytics Data on IT Professionals (July 13, 2021). University of Chicago, Becker Friedman Institute for Economics Working Work No. 2021-56, Available at SSRN: https://ssrn.com/abstract=3843197 or http://dx.doi.org

Yuksel D, McKee GB, Perrin PB, et al. Sleeping when the world locks down: Correlates of sleep health during the COVID-19 pandemic across 59 countries. Sleep Health. 2021;7(2):134-142. doi:10.1016/j.sleh.2020.12.008

Conroy DA, Hadler NL, Cho E, et al. The effects of COVID-19 stay-at-home order on sleep, health, and working patterns: a survey study of US health workers. J Clin Sleep Med. 2021;17(2):185–191.

Morse KF, Fine PA and Friedlander KJ (2021) Creativity and Leisure During COVID-19: Examining the Relationship Between Leisure Activities, Motivations, and Psychological Well-Being. Front. Psychol. 12:609967. doi: 10.3389/fpsyg.2021.609967

Morelli M, Cattelino E, Baiocco R, Trumello C, Babore A, Candelori C and Chirumbolo A (2020) Parents and Children During the COVID-19 Lockdown: The Influence of Parenting Distress and Parenting Self-Efficacy on Children's Emotional Well-Being. Front. Psychol. 11:584645. doi: 10.3389/fpsyg.2020.584645

Rehman U, Shahnawaz MG, Khan NH, et al. Depression, Anxiety and Stress Among Indians in Times of Covid-19 Lockdown. Community Ment Health J. 2021;57(1):42-48. doi:10.1007/s10597-020-00664-x

De' R, Pandey N, Pal A. Impact of digital surge during Covid-19 pandemic: A viewpoint on research and practice. Int J Inf Manage. 2020;55:102171. doi:10.1016/j.ijinfomgt.2020.102171

Peters, P, Tijdens, KG and Wetzels, C (2004), 'Employees' opportunities, preferences, and practices in telecommuting adoption', Information & management, 41(4), pp. 469–482.

Elise Gould, Melat Kassa, Young workers hit hard by the COVID-19 economy, Economic Policy Institute: October 14, 2020. https://www.epi.org/publication/young-workers-covid-recession/

Scikit-learn: Machine Learning in Python, Pedregosa et al., JMLR 12, pp. 2825-2830, 2011.

G. King and L. Zeng, ''Logistic regression in rare events data,'' Political Anal., vol. 9, no. 2, pp. 137–163, 2001

A. Liaw and M. Wiener, ''Classification and regression by randomForest,'' R News, vol. 2, pp. 18–22, Dec. 2002.

Webb G.I. (2011) Naïve Bayes. In: Sammut C., Webb G.I. (eds) Encyclopedia of Machine Learning. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-30164-8_576

KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data MiningAugust 2016 Pages 785–794https://doi.org/10.1145/2939672.2939785

The Journal of Machine Learning ResearchVolume 712/1/2006 pp 551–585

Smola, A.J., Schölkopf, B. A tutorial on support vector regression. Statistics and Computing 14, 199–222 (2004). https://doi.org/10.1023/B:STCO.0000035301.49549.88

KEY TERMS AND DEFINITIONS

t - calculated difference in units of standard error.

df – the largest number of logically independent values, that can vary within the dataset.

p - probability that the sample data results occurred by chance.

H0 – Null Hypothesis, the true difference between the group means is zero.

HA – Alternate Hypothesis, the true difference between the group means is not zero.

CI – Confidence Interval, it measures the degree of certainty in a sampling method. The most common probability limit is 95%, which is also the limit specified in this work.

Accuracy - The number of correct prediction divided by the total number of predictions.

F1 Score - harmonic mean of the recall and precision.

Recall - number of true positives divided by the sum of the number of true positives and number of false negatives.

Precision - number of true positives divided by sum of the number of true positives and the number of false positives.

APPENDIX 1

Questionnaire Items
i. Age (12-18, 19-25, 25-32, 33-40, 41-50, 51-60)
ii. Gender (Male, Female, Other)
iii. Occupation (Working Professional, Student in School, Student in College, Homemaker, Entrepreneur)
iv. Daily average work hours before remote work enforcement
v. Daily average work hours after remote work enforcement
vi. Daily average time spent travelling

(Questions **vii.** To **xv.** are answered with the options: Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree)
vii. My Productivity levels have increased working remote
viii. My sleep cycle is a lot more balanced with remote work
ix. I have been able to find time to pick up a hobby
x. I spend more time with family and better connect with them
xi. My stress levels have reduced working remote
xii. I have been able to commit more time to self-care
xiii. It is easy to adjust to a purely online setting to work
xiv. My home environment allows me to focus on work

(Questions **xv.** to **xxii.** are answered with options: Yes, No)
xv. Zero travel time to my work appeals to me
xvi. I like being close to my loved ones all day
xvii. I prefer not having to do much physical work
xviii. I like not having to interact with my colleagues in person
xix. I dislike not being able to go outside and exercise
xx. I don't like not being able to travel
xxi. I dislike the effect that remote work has on my social life
xxii. Overall, I prefer: (Remote Work, Work in-person)