Computers & Security 149 (2025) 104204

Contents lists available at ScienceDirect

Computers & Security

journal homepage: www.elsevier.com/locate/cose

Full Length Article

Explanatory and predictive modeling of cybersecurity behaviors using protection motivation theory

Uzma Kiran [a], Naurin Farooq Khan [a], Hajra Murtaza [a], Ali Farooq [b,*], Henri Pirkkalainen [c]

[a] Riphah International University, Islamabad, Pakistan
[b] University of Strathclyde, Glasgow, United Kingdom
[c] Tampere University, Tampere, Finland

ARTICLE INFO

Keywords:
Machine learning
Predictive modeling
Explanatory modeling
Protection motivation theory
Smartphone security behavior
Cybersecurity behavior

ABSTRACT

Context: Protection motivation theory (PMT) is the most frequently used theory in understanding cyber security behaviors. However, most studies have used a cross-sectional design with symmetrical analysis techniques such as structure equation modeling (SEM) and regression. A data-driven approach, such as predictive modeling, is lacking and can potentially evaluate and validate the predictive power of PMT for cybersecurity behaviors.
Objective: The objective of this study is to assess the explanatory and predictive power of PMT for cyber security behaviors related to computers and smartphone.
Method: An online survey was employed to collect data from 1027 participants. The relationship of security behaviors with *threat appraisal (severity and vulnerability)* and *coping appraisal (response efficacy, self-efficacy and response cost)* components were tested via explanatory and predictive modeling. Explanatory modeling was employed via SEM, whereas three machine learning algorithms, namely Decision Tree (DT), Support Vector Machine (SVM), and K Nearest Neighbor (KNN) were used for predictive modeling. Wrapper feature selection was employed to understand the most important factors of PMT in predictive modeling.
Results: The results revealed that the *threat severity* from the *threat appraisal* component of PMT significantly influenced computer security and smartphone security behaviors. From the *coping appraisal, response efficacy* and *self-efficacy* significantly influenced computer and smartphone security behaviors. The ML analysis showed that the highest predictive power of PMT for computer security was 76 % and for smartphone security 68 % by KNN algorithm. The wrapper feature selection approach revealed that *the most important features in predicting security behaviors are self-efficacy, response efficacy and intention to secure devices*. Thus, the findings indicate the complementarity of the cross-sectional and data driven methods.

1. Introduction

Cybersecurity has been an important area of research for the last two decades. With the notion of humans as the weakest link in cybersecurity, behavioral aspects have also been given due research diligence recently (Schneier, 2015). Several studies have highlighted the impact of human behavior in securing cyberspace. Various theoretical constructs and frameworks from fields such as social sciences, information systems, and psychology have been used to understand the influence of socio-technical variables on varied types of cybersecurity behaviors in organizational setups, such as policy compliance and non-compliance (Moody et al., 2018), and home-users contexts, such as general and specific security behaviors (Khan et al., 2022; Y. Li et al., 2021). Notwithstanding the context, one of the most frequently used theories in understanding cybersecurity behavior is Roger's protection motivation theory (PMT) (Rogers, 1975) and has been reported consistently to be the most frequent by many systematic reviews (R. E. Crossler et al., 2013; Haag et al., 2021; Mou et al., 2017; Siponen et al., 2024). While PMT has been used to understand user's security behaviors in both organizational and home, it has been found more appropriate in understanding home-users motivations and behaviors in comparison to organizational ones because it deals with individual behaviors, and personal perceptions of risks and motivations instead of those related to the organization (Sommestad et al., 2015a). The individuals in home-user contexts can be students, children, adults and senior citizens (Farooq et al., 2015; Y. Li et al., 2021) who interact with cyberspace

* Corresponding author.
E-mail addresses: ali.farooq@strath.ac.uk, afk@ieee.org (A. Farooq).

https://doi.org/10.1016/j.cose.2024.104204
Received 9 July 2024; Received in revised form 10 October 2024; Accepted 7 November 2024
Available online 9 November 2024

devoid of organizational policies, rules and regulations, and their protective behavior is understood better by *threat* and *coping appraisal* as per PMT. Other the other hand, employees' or adult workers' security behaviors are govern by organizational policies, rules, and regulations in addition to their individual threat and *coping* appraisal (Sommestad et al., 2015a).

Studies addressing the effects of PMT's constructs on cyber security behaviors are largely based on cross-sectional research design and symmetrical analysis techniques such as regression (multiple, hierarchical) and structural equation modeling (SEM), for example, (Farooq et al., 2019; Haag et al., 2021; Mou et al., 2017). At one end, the cross-section design inhibits researchers from examining the questions that implicitly or explicitly point toward temporal causality (Maier et al., 2023). On the other, both analysis techniques, which are explanatory in nature (Forster and Sober, 1994; Shmueli, 2010), aim to test the hypotheses and report on the strengths of interactions between the threat and *coping appraisal*s of PMT and cybersecurity behavior. The premise of such modeling is based on the compensatory nature of interactions where the shortfall of one antecedent's influence in the PMT model is compensated by others' influence (Alwabel and Zeng, 2021). However, regression and SEM assume a linear relationship between antecedents of cybersecurity behavior (Alwabel and Zeng, 2021), which sets both empirical and theoretical limitations to understanding the phenomenon.

To deal with the limitations of both research design and analysis methods, it is essential to consider alternative methods for understanding cybersecurity behaviors, for example, Maier et al. (2023) recommended studying casualty beyond symmetrical modeling. Existing information systems research proposes configurational and asymmetrical modeling techniques such as fuzzy-set qualitative comparative analysis (fsQCA) (Mithas et al., 2022), and neural networks (NN) (Tarhini et al., 2024) to overcome the limitations of symmetrical or explanatory modeling. The fsQCA approach complements the statistical modeling approaches by identifying sufficient causal configurations i.e. combination of factors for given behaviors (Dahabiyeh et al., 2023), or the sequence (Sun et al., 2020). fsQCA is helpful where the relationships are complex and multiple conditions hold with configurational dependencies. fsQCA has been known to better address the problems related to small samples and qualitatively complements SEM by providing particular configurations and combinations. NN has been found better at studying non-linear relationships and gives greater predictive modeling power (Liébana-Cabanillas and Lara-Rubio, 2017; Shmueli, 2010). Unlike regression and SEM, this data-driven modeling seeks to identify the predictive power of the model and does not consider the correlations between the antecedents of the model (Alwabel and Zeng, 2021). Predictive modeling is specifically important for advancing and validating theories and for evaluating the predictive accuracy of models (Shmueli and Koppius, 2011). The use of non-symmetrical and predictive modeling with explanatory modeling is evident in other domains, such as the adoption of technology (Alwabel and Zeng, 2021; Bahari et al., 2023), environment (Almheiri et al., 2024) and financial sustainability (Arpaci, 2023) and education (Alshurideh et al., 2023; Tarhini et al., 2024).

Despite their utilities, the configurational and predictive modeling in cybersecurity behavioral research is almost non-existent. Recent studies on cybersecurity behaviors have called for a better understanding of the interplay of explanatory and predictive modeling (Alassaf and Alkhalifah, 2021). Advancing the understanding of predictive modeling will not only diversify analysis techniques to overcome biases of cross-sectional research designs (Maier et al., 2023) but also will help with a better understanding of machine learning algorithms (Khan et al., 2021). The predictive power of the explanatory models ascertains the quantity of actual explanations (Maier et al., 2023). Moreover, the completeness of the explanatory models has been assessed via predictive models as benchmarks (Fudenberg et al., 2019) and sets precedence for complete explanations.

This study addresses the gaps identified in the use of PMT in cybersecurity behavioral research by diversifying analysis techniques to deal with both design and limitations of existing analysis methods (Alassaf and Alkhalifah, 2021; Maier et al., 2023). The overreaching aim of this study is *to compare the explanatory and predictive modeling approaches to explain cybersecurity behaviors*. This allows us to bridge the research gaps by augmenting SEM analysis with data-driven and machine-learning-based predictive analysis. This study compares the two different methods (explanatory vs. predictive) in the contexts of computer and smartphone security behaviors that have been previously indicated as two critical yet distinctive IT-specific settings to cybersecurity behaviors (Thompson et al., 2017). Specifically, this study addresses cybersecurity behavior in the home-user context to account for the predictive accuracy of the PMT on cybersecurity behavior. The explanatory modeling was accomplished with covariance-based structural equal modeling (CB-SEM) to gain insights into the explanatory power of PMT for computer security behavior (CSB) (Ng et al., 2009; Thompson et al., 2017) and smartphone security behavior (SSB) (Zhou et al., 2020). Then, a combination of linear and non-linear machine learning algorithms (ML) was used to understand the predictive power of PMT and to assess the protection motivation for cybersecurity behaviors. As a contribution, the study indicates the complementarity of the explanatory and predictive approaches to indicate similar findings and the distinctiveness of the predictive modeling to indicate more nuanced effects. This serves as a key step for Information Systems research to seek the potential of predictive modeling in behavioral research.

The paper is organized as follows. Section 2 presents the background and literature review on cybersecurity behavioral research in terms of computer security and smartphone security under the lens of PMT. The explanatory modeling and predictive modeling are explained in Subsections 2.2 and 2.3 followed by key differences between the two in Subsection 2.4. Section 3 reports on the details of the proposed models and hypothesis development. The details about the research methodology such as instruments used, sampling strategy and data collection procedure, are given in Section 4 and its subsequent Subsections. The results of explanatory and predictive modeling are discussed in Section 5 followed by discussion. The conclusion and future work is given in Section 6.

2. Background and literature review

2.1. Protection motivation theory

Protection Motivation Theory (PMT) was developed by Rogers (Rogers, 1975) in 1975 to help explain an individual's engagement in protective behaviors by focusing on fear appeals. Later, the theory was revised to include cognitive factors for understanding protective behaviors (Rogers, 1983). PMT can and has been used in various situations involving any threat (Rogers and Prentice-Dunn, 1997). PMT has been used in health, medical, social, personal and economic domains (Farooq et al., 2020; Floyd et al., 2000). According to PMT, the manifestation of protective behavior is the outcome of cost-benefit analysis; the individual compares risks and costs and decides whether to take protective actions to eliminate the risk (Rogers, 1983; Sommestad et al., 2015a). The cost-benefit analysis is carried out using two components of PMT: *threat appraisal* and *coping appraisal*. The *threat appraisal* appraises how likely an unwanted threat results in consequences while *coping appraisal* is the engagement in a protective behavior via the tradeoff between the effectiveness of the coping response and its costs (Rogers, 1983). This means that an individual should perceive that one is vulnerable to a threat and that the consequences of the threat are severe. At the same time, the individual should perceive that one can enact protective behavior that is effective for mitigating threats, and the cost of enactment is compensated with its benefits (Rogers, 1983). Both *threat* and *coping appraisals* create motivation (intention) among the individuals,

further eliciting actual behavior. *Threat appraisal* is operationalized using two constructs: *threat severity* and *threat vulnerability*. The former is the perception of threat impact when it becomes real, and the latter is the likelihood of falling prey to a threat. A higher severity and vulnerability would create protection motivation and, thus behavior. *Coping appraisal* consists of three constructs: *response efficacy, self-efficacy* and *response cost. Response efficacy,* is one's belief that an adaptive response is effective, *self-efficacy* is one's belief about ability to successfully perform an adaptive response, and *response cost* is the psychological or physical cost associated with adopting a control measure. Fig. 1 shows the factors involved in PMT.

2.2. PMT and cybersecurity behavior

Almost two decades of behavioral cybersecurity research have repeatedly identified PMT as the most frequently used theory in understanding cybersecurity behaviors (Haag et al., 2021; Khan et al., 2022; Lebek et al., 2014; Mou et al., 2017, 2022). The theory has been touted as an appropriate theory for studying cybersecurity behavior, with its components mapping well to the security concepts (Sommestad et al., 2015b).

In the cybersecurity behavior context, *threat vulnerability* has been defined as the likelihood of occurrence of a threat and that an individual is likely to be exposed to the threat (Haag et al., 2021; Johnston and Warkentin, 2010). *Threat severity* has been defined as the severity of the negative consequences that can occur due to the manifestation of that threat (Haag et al., 2021). It is the severity of the repercussions when a threat materializes. If an individual perceives that one is vulnerable to computer security or smartphone threats and that the consequences of such threats are harmful to him, it results in one's adoption of computer and smartphone security behaviors. *Coping appraisal* is the cognitive process that allows individuals to engage in cybersecurity behaviors (Doane et al., 2016). *Self-efficacy* is the skills of an individual to enact a cybersecurity practice (Haag et al., 2021). *Response efficacy* in cybersecurity context is the perception that secure behavior benefits the individual (Crossler and Bélanger, 2014) by mitigating cyber threats. *Response cost* is defined as any cost associated with performing cybersecurity behavior (Haag et al., 2021; Mou et al., 2017). The nature of cost can be financial or temporal, and it can also manifest in the effort required by an individual or inconvenience while performing a cybersecurity response. Individuals carry out a calculated decision to relinquish a cybersecurity response if the cost is higher than the severity of the cyber threat (R. Crossler and Bélanger, 2014).

2.3. Related work

Here, we report the latest studies that use PMT by consulting the latest systematic literature reviews and individual studies published at the time of the writing. Specifically, we utilized multiple databases (e.g., Google Scholar, Scopus, IEEE Xplore) to identify the literature, identify the relevant works, and gain a comprehensive understanding of the topic by going back and forth between the studies. The summary of related work is provided in Table 1.

A study by (Sharma and Aparicio, 2022) has extended protection motivation theory to include organizational and team culture to study their influence on threat and *coping appraisal* of PMT. They showed that *threat appraisal* and *coping appraisals* (except *response cost*) were influenced by organizational and team cultures, which, in turn, explained the motivation to comply with organizational security policies. Another study (Ogbanufe et al., 2023) tested the PMT model to explain the intention to violate organizational security policies in three ways: 1) as a standalone PMT model, 2) as an integrated PMT model with stewardship theory, and 3) as a stand-along stewardship model. The results showed that PMT alone was able to explain the intention to violate security policies better than the stewardship model; however, the integrated model was slightly better. The *coping appraisal* (*self-efficacy*, response efficacy and *response cost*) had a significant association, while threat severity did not show a significant association. A recent study examining entrepreneurs' security behavior employed PMT and extended it with subjective norms, threat awareness and affective response (Luuk et al., 2023). The results showed that entrepreneurs with high threat severity and vulnerability perceptions had a higher intention to protect against ransom-ware, however, their perception of *self-efficacy* and the *response efficacy* had a negative significant influence on protection motivation. Similarly, another PMT model (L. Li et al., 2022) was extended with organizational effort and employees' cybersecurity awareness; the two constructs were taken from organizational effort theory and the theory of planned behavior (TPB). The results of the study revealed that the

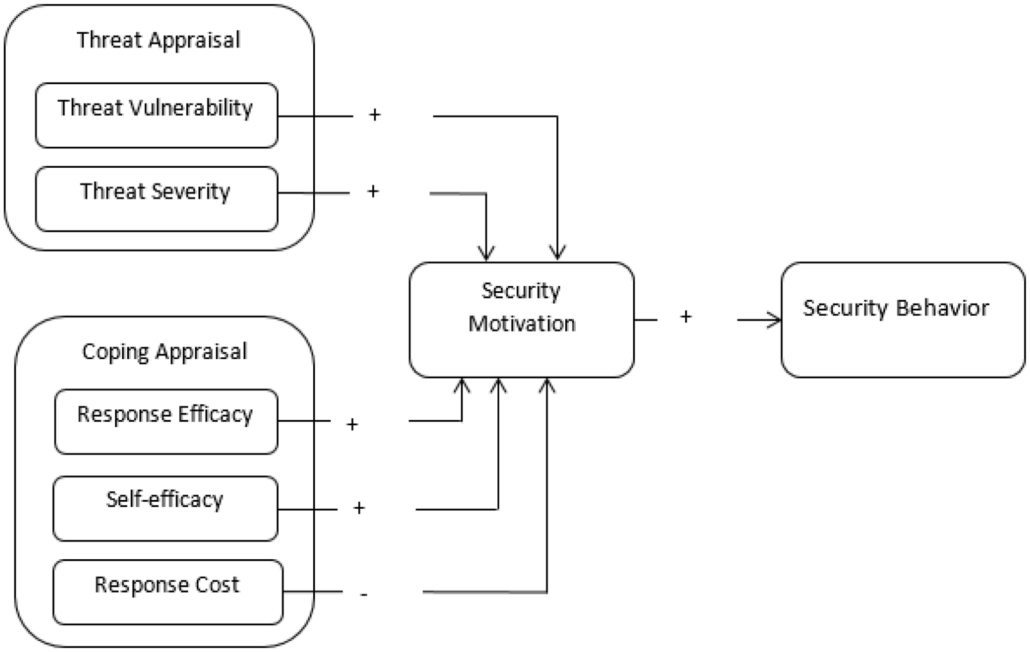

Fig. 1. Protection motivation theory in general.

Table 1
Summary of the literature.

Study	Context	Sample size	Country	Research Design	Analysis Technique
(Sharma and Aparicio, 2022)	IT employees	341	USA	Cross sectional	PLS-SEM
(Ogbanufe et al., 2023)	Employees	339	USA	Cross sectional	PLS-SEM
(Luuk et al., 2023)	Entrepreneurs	1020	Netherlands	Cross sectional	SEM
(L. Li et al., 2022)	Employees	387	USA	Cross sectional	SEM
(Rajab and Eydgahi, 2019)	Employees	206	USA	Cross sectional	PLS-SEM
(Dang-Pham and Pittayachawan, 2015)	Students	252	Australia	Cross sectional	Bayesian SEM
(Vrhovec and Mihelič, 2021)	University Faculty	255	Slovenia	Cross sectional	SEM
(Mills and Sahi, 2019)	Home users	72	Not specified	Cross sectional	PLS-SEM
(Thompson et al., 2017)	General users	629	USA	Cross sectional	PLS-SEM
(Tsai et al., 2016)	Employees	988	USA	Cross sectional	Regression
(Farooq et al., 2019)	Students	125	Kenya	Cross sectional	SEM
Smartphone Security Studies					
(R. E. Crossler et al., 2014)	Employees and students	444		Cross sectional	PLS-SEM
(Dang-Pham and Pittayachawan, 2015)	Students	252	Australia	Cross sectional	Bayesian SEM
(Thompson et al., 2017)	General users	629	USA	Cross sectional	PLS-SEM
(Hovav and Putri, 2016)	Employees	230	Indonesia	Cross sectional	SEM
(Tu et al., 2019)	Employees	122	America, Europe, Asia	Cross sectional	PLS-SEM
(Verkijika, 2018)	General users	428	South Africa	Cross sectional	PLS-SEM
(Giwah et al., 2019)	General users	390	USA	Cross sectional	PLS-SEM
(Knapova et al., 2021)	General users	502	Czech Republic	Cross sectional	Regression
(Ameen et al., 2021)	Employees	1735	USA, UK and UAE	Cross sectional	PLS-SEM

threat appraisal (*threat severity* and vulnerability) had no significant influence on protective behavior while *response efficacy* and *self-efficacy* had a positive while *response cost* had a negative significant influence. PMT was integrated with TPB, general deterrence theory (GDT) and organizational theory to understand cybersecurity compliance's influencing factors in higher education institutes (Rajab and Eydgahi, 2019). The study found that PMT constructs were the best-explaining factors with perceived vulnerability, and response efficacy to have significant positive while response cost to have a negative significant effect on cybersecurity compliance. A survey conducted in Slovenian universities employed PMT to understand protection motivation and differentiated between *threat appraisal* towards an individual and an organization (Vrhovec and Mihelič, 2021). The results depicted that threat severity and threat vulnerability had an indirect effect on protection motivation mediated by fear. In *coping appraisal, response efficacy* had a significant positive effect on employees' motivation to protect themselves while *self-efficacy*'s positive relationship was dampened by the fear of cyber-attacks.

There are a number of studies examining smartphone security behavior, and bring your own device (BYOD) in home contexts (Brodin and Rose, 2020; Butler, 2020; Palanisamy et al., 2020). Many researchers have found low compliance towards the BYOD security policies in an organizational context while others have descriptively reported the low state of smartphone security practices among general users (Breitinger et al., 2020; Chin et al., 2020; Das and Khan, 2016; Harris et al., 2014; Jones and Chin, 2015; Khan et al., 2023, 2022; Mai and Tick, 2021; Nowrin and Bawden, 2018; Shah and Agarwal, 2020; Stylios et al., 2016; Zhang et al., 2017). Although the number of smartphone security studies employing theoretical underpinning is scarce, PMT has been reported to be one of the top theories employed to study smartphone security behavior (Palanisamy et al., 2020) (Dawie et al., 2022).

One of the earliest studies was by (R. E. Crossler et al., 2014), in which factors affecting the BYOD policy compliance were studied. The results revealed that *self-efficacy* and *response efficacy* were the dominant factors in explaining the motivation to comply with BYOD policies. Another study (Dang-Pham and Pittayachawan, 2015) was carried out to understand the malware avoidance behavior of the students and their compliance with the an Australian university's BYOD policies. The study found that all the protection motivation elements had a significant influence on smartphone security intention in university settings though the differentiation of the effects sizes was not carried out. The PMT constructs were extended to include social influence and psychological ownership by (Thompson et al., 2017), in which it was found that perceived vulnerability, self-efficacy and response cost have a significant influence on intention to comply with smartphone security behavioral intention. Similarly, Verkijika (2018) augmented the PMT model with anticipated regrets to understand the smartphone security behaviors of users in South Africa. The results revealed that *self-efficacy* had a direct effect on smartphone security intention while perceived vulnerability and severity were mediated by anticipated regret to explain the smartphone security intention and behavior. Another survey was conducted on 230 employees of an Indonesian organization by blending PMT with organizational justice theory (Hovav and Putri, 2016). The study found that *response efficacy* and justice had a strong influence on employees' intention towards BYOD policy compliance. On the other hand (Tu et al., 2019) leveraged PMT to understand the key factors responsible for employees' compliance with BYOD policies. The results showed *self-efficacy*, perceived vulnerability, perceived severity and *response efficacy* to have a significant positive effect while *response cost* to have a significant negative influence on smartphone security intention. A recent study (Knapova et al., 2021) employed the health belief model along with PMT to understand the smartphone security determinants. The study employed a total of 331 participants from the Czech reported a positive influence of perceived severity and *self-efficacy* on smartphone security behavior along with other factors such as security orientation and personal experience with digital threats. A recent study (Ameen et al., 2021) made use of PMT along with general deterrence theory (GDT) and theory of reasoned action to understand the determinants of smartphone security more holistically by drawing samples from three countries (UK, USA and UAE). The results suggested mixed findings for UK, USA and UAE due to cultural differences.

The literature on adoption of protection motivation theory (Table 1) in cybersecurity domain shows that most of the research done is done by symmetric analysis i.e. explanatory modeling. Whereas the use of predictive modeling is almost nonexistent.

2.4. Explanatory and predictive modeling

Explanatory modeling is the use of statistical models that tests causal hypotheses. Both regression models and structural equation models fall under this, and they rely on observational data. The hypotheses are either tested based on association to test the causality or the strength of relationships R^2 (Byrne, 2013). In regression models, association between the independent and dependent variables by understanding the change occurred in independent variables also changes the dependent

variables (Panovska-Griffiths et al., 2021). It is achieved by fitting the best fit line and seeing the dispersion of data around it. Structural equation modeling is a multivariate technique that evaluates multivariate causal relationships (Fan et al., 2016). In contrast to regression, SEM can show direct as well as indirect effects on causal relationship among variables. SEM does this by combining confirmatory factor analysis with path analysis (Fan et al., 2016). Both regression and SEM are widely used as a statistical modeling in behavioral research.

Predictive modeling is based on machine learning, which is a subset of artificial intelligence (AI). Different mathematical based algorithms are employed on the given data to learn and formulate an understanding of a given phenomenon (Murphy, 2012). The goal of ML is to predict the target output based on a selected number of features hence it provides empirical evidence via data driven methodology (Alwabel and Zeng, 2021). The prediction can be done either by supervised, unsupervised or reinforcement learning. In supervised learning, the target output is already known (Murphy, 2012). In unsupervised learning, the target output is not known and the focus is on organization of the data based on similarities known as cluster (Alwabel and Zeng, 2021). Whereas in reinforcement learning; the focus is developing a solution via hit and trial that is achieved by doing multiple iterations. A number of different types of machine learning algorithms are present to predict the target outcome and they are either linear, non-linear or hybrid in nature (Alwabel and Zeng, 2021). The linear algorithms assumes a linear relationship between the input and output variables while the non-linear algorithms are able to capture complex relationships between the two foregoing the linear relationship. They hybrid ML algorithms on the other hand, leverage the simplicity of linear algorithms with the flexibility of non-linear ones. The algorithms work by computing the mean difference of unobserved data and predicted data and measure the predictive power of the model (Murphy, 2012).

2.5. Differences in explanatory and predictive modeling approaches

Explanatory modeling provides information about the direction and strength of the relationship between independent and dependent variables (Forster and Sober, 1994), and statistical techniques are used for testing causal theory (Shmueli, 2010). It helps to understand the underlying structure of data and can identify if a causal relationship exists between the variables (Shmueli and Koppius, 2011). The explanatory modeling however, takes into account assumptions about the data that should be met before analysis is carried out. These assumptions are that the variables should be normally distributed, the relationship between the variables should be linear, the model should have unique information for parameter estimation, and the measurement should be devoid of errors. These assumptions limit the explanatory analysis to accurately estimate the relationships between variables, due to which the results can be inaccurate or biased. Predictive modeling is the application of the algorithm to data to predict future observations given the historical data (Geisser, 1993). It does not consider the same assumptions required in explanatory modeling and is trained on a wide range of data (Table 2). Prediction allows for the identification of patterns and non-linear relationships that may not be brought to light by explanatory modeling (Shmueli, 2010). It does not necessarily require a theoretical foundation, thereby treating the model as a black box (Shmueli and Koppius, 2011). The understanding of the phenomenon is then gained by the data which is collected from the environment (Alwabel and Zeng, 2021).

The field of behavioral cybersecurity is dominated by explanatory models, while predictive models are poorly understood and not given their due place (Alassaf and Alkhalifah, 2021; Khan et al., 2022). Predictive modeling offers improvements to existing explanatory models by capturing complex relationships and pattern (Shmueli, 2010). It should be noted that both explanatory and predictive modeling play various roles in the generation and testing of theories (Shmueli and Koppius, 2011). In fact, it bridges the gap between theory and practice by offering predictive power which is different from the explanatory power (Forster and Sober, 1994). By quantifying the predictability level of the phenomenon of interest, it creates predictive accuracy benchmarks. Since the predictive power is more accurate in prediction than explanation, adding predictive analysis can improve the accuracy and reliability of a causal inference (Alwabel and Zeng, 2021). Hence combination of prediction and explanation has multiple benefits. Prediction can be used to test the validity of results derived from explanatory modeling. This can be done by measuring the accuracy of the predicted target outcome; the greater the accuracy, the more confidence in the validity of explanatory modeling. Researchers can improve the reliability of the causal inference by adding predictive analysis. Apart from its practical usefulness, it can be used in building and testing theories (Dubin, 1969).

Table 2
Differences in explanation and prediction.

Consideration	Explanation	Prediction
Main Purpose	Used for understanding the relationship between independent variables and dependent variable.	Used for prediction of future outcomes by learning on historical data irrespective of the distribution or relationship between the variables
Mechanism	Takes into consideration the co-relation between different variables.	Computes the average difference between the unknown data and predicted data
Black box vs White box	Takes into consideration the verification of theoretical assumptions after data is collected, modelled and verified.	Takes into consideration the relationship between variables as a black box which may be conceived by the environment
Deductive vs inductive	Heavy reliance on theory.	Reliance on data
Underlying assumptions	Makes several assumptions such as multivariate normal, linearity of model, model identification and free of error.	Does not make assumptions about the data and the model

3. Hypothesis development and research models

In this study, we propose two models based on PMT to understand its explanatory and predictive power for computer security and smartphone security. Below, we describe hypotheses related to computer security and smartphone security and the two research models (Fig. 2) that will be further examined with the explanatory and predictive methods.

3.1. Threat appraisal

Previous literature on computer security has found a significant positive association of *threat vulnerability* and intention to secure computer devices (Hina et al., 2019). Similarly, *threat severity* has been found to be significant in influencing the intention to secure computer devices (Ifinedo, 2012; Posey et al., 2011; Vance et al., 2012). The hypothesis related to *threat appraisal* in computer security behavior (CSB) are described below (H1- H2)

- H1a: *Perceived threat vulnerability* has a positive influence on intention to secure computer device.
- H2a: *Perceived threat severity* has a positive influence on intention to secure computer device.

The positive association of *threat vulnerability* on intention to secure smartphone security devices has been reported by a number studies (Dang-Pham and Pittayachawan, 2015; Thompson et al., 2017; Verkijika, 2018). Previous literature on smartphone (Hovav and Putri, 2016; Knapova et al., 2021; Thompson et al., 2017; Verkijika, 2018) has reported a significant positive relationship between *threat severity* and smartphone security intention. The hypotheses H1b-H2b are related to

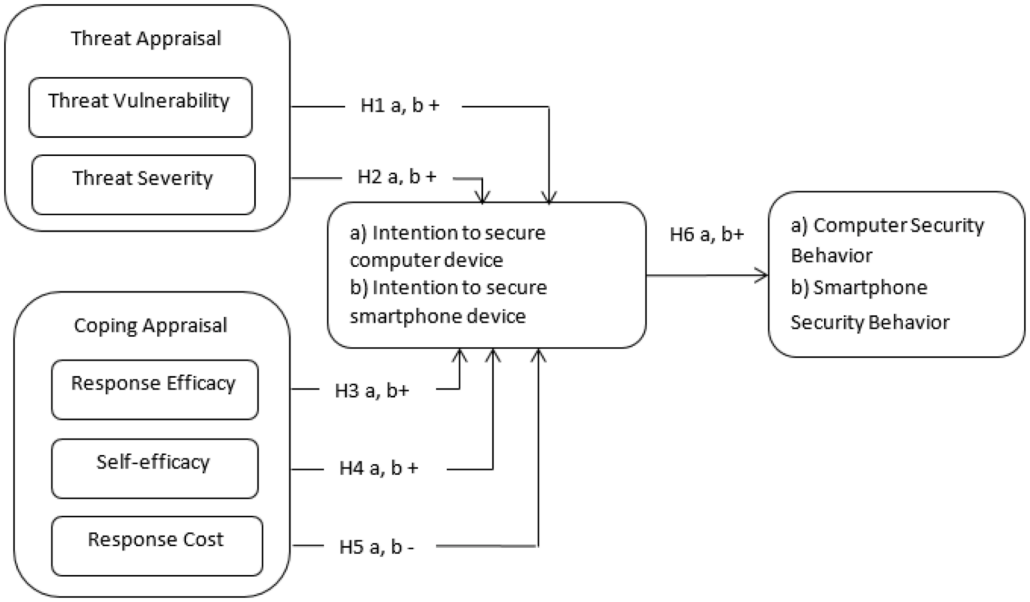

Fig. 2. PMT-based computer security and smartphone security models to be examined with explanatory and predictive approaches.

threat appraisal in smartphone security (SSB).

- H1b: *Perceived threat vulnerability* has a positive influence on intention to secure smartphone device.
- H2b: *Perceived threat severity* has a positive influence on intention to secure smartphone device.

3.2. Coping appraisal

Self-efficacy has been regarded as a highly significant predictor of protective intent (L. Li et al., 2019), the best measure of intent (Posey et al., 2015) and the most direct determinant (Johnston and Warkentin, 2010). *Response efficacy* has been found to have a comparatively stronger effect on the protection motivation than the *threat appraisal* dimension of PMT (Posey et al., 2015) and positive association with intention to comply with (Johnston and Warkentin, 2010). Various studies have found a negative influence of *response cost* on cybersecurity practices (Posey et al., 2015; Tsai et al., 2016; Vance et al., 2012).

The hypotheses H3(a) – H5(a) are related to *coping appraisal* in computer security.

- H3a: Perceived *response efficacy* associated with enactment of cybersecurity coping mechanisms has a positive influence on intention to secure computer device.
- H4a: Perceived *self-efficacy* associated with enactment of cybersecurity coping mechanisms has a positive influence on intention to secure computer device.
- H5a: Perceived *response cost* (RC) associated with cybersecurity coping mechanism has a negative influence on intention to secure computer device.

Smartphone *Self-efficacy* has been positively associated with the intention to secure smartphone device in (Dang-Pham and Pittayachawan, 2015; Giwah et al., 2019; Knapova et al., 2021; Thompson et al., 2017; Verkijika, 2018). For *response efficacy* in smartphone security, studies have found (Dang-Pham and Pittayachawan, 2015; Giwah et al., 2019) significant positive associations between the two. The literature on smartphone security has reported a negative impact of *response cost* on the intention to secure smartphone devices (Dang-Pham and Pittayachawan, 2015; Hovav and Putri, 2016; Thompson et al., 2017; Tu et al., 2019; Verkijika, 2018). The following hypotheses are posed for smartphone security;

- H3b: Perceived *response efficacy* associated with enactment of smartphone security coping mechanism has a positive influence on intention to secure smartphone device.
- H4b: Perceived *self-efficacy* associated with enactment of smartphone security coping mechanism has a positive influence on intention to secure smartphone device.
- H5b: Perceived *response cost* associated with smartphone security coping mechanism has a negative influence on intention to secure smartphone device.

3.3. From intention to behavior

PMT takes into account the security behavior that is mediated by security intention. The premise is that individuals' intentions to secure their devices will eventually be translated into behaviors. Studies have found a significant positive influence of cybersecurity intentions on computer and smartphone security behaviors (Giwah et al., 2019; Thompson et al., 2017; Verkijika, 2018). The hypotheses H6 (a) and H6 (b) are related to computer security and smartphone security behaviors respectively.

- H6a: The intention to secure computer device has a significant positive influence on computer security behavior.
- H6b: The intention to secure smartphone device has a significant positive influence on smartphone security behavior.

4. Research methodology

A quantitative methodology was employed to carry out this research. The next Subsections discuss the instruments employed, sampling, and data collection procedure.

4.1. Measures

The measures used in this study are adopted from previous literature. The instruments' items or their response types were not changed to ensure their reliability and validity. This study makes use of the security behavior intention scale (SeBIS) scale to capture computer security behavior which was adopted from a previous study (Egelman et al.,

2016; Egelman and Peer, 2015). SeBIS is a validated scale consisting of four dimensions and has 16 items. These four dimensions are: Device Security (DS), Password Generation (PG), Proactive Awareness (PA) and Updating (UP). Although SeBIS had a measurement scale more suitable for behavior (never to always), the authors suggested it as an intention scale as real behaviors are hard to record through a self-reported design. However, the scale has been found to correlate significantly with actual behavior measured objectively and should be considered sufficient for measuring self-reported behavior (Egelman et al., 2016). It measures device securement, updating behavior, password protection and proactive awareness. The items are measured on a 5-point Likert scale with 1 as 'Never' and 5 as 'Always'. It has a high internal consistency with Cronbach's alpha = 0.81 in the original study. It also has an established criterion validity, showing high correlations with actual security behaviors.

The scale employed for smartphone security was adopted from a previous study (Huang et al., 2020). It is also a 5-point Likert scale, with responses measuring from 1 as Never to 5 as always. It consists of two dimensions, namely technical security and social security. The scale has been developed by item collection and expert evaluation by carrying out exploratory factor analysis and confirmatory factor analysis. The scale has been validated for its convergent validity by correlating it with the SeBIS scale. The Cronbach's alpha = 0.79 shows internal consistency and shows that it is a reliable instrument in the original study.

The PMT elements were adopted from previous studies (Thompson et al., 2017). *Self-efficacy, response efficacy, response cost, threat vulnerability*, and *threat severity* are measured on a 7-point Likert scale (1 = strongly disagree to 7 = strongly agree). The responses range from 'strongly disagree' coded as 1, to 'strongly agree' coded as 7, with 'neutral' in the middle coded as 4. For *self-efficacy, response efficacy, threat vulnerability*, and *threat severity*, a higher score depicts a higher perception of the *coping appraisal* and *threat appraisal* respectively. For *response cost*, high scores mean a low perception of *coping appraisal*. The motivation to protect one's devices is also taken from the previous studies which is measured as security intention. The items for PMT elements are given in detail in Appendix 1.

4.2. Data collection

The data for the study was collected via survey. We designed the survey in Google Forms to measure protection motivation, computer security and smartphone security behaviors. Since the focus was on the home-users, we opted to collect data from university students who represent a sub-set of home-users. The survey was carried out in Pakistan. We opted for a diverse sample (Table 3) by recruiting students from multiple universities. For this purpose, the first author of the study contacted several instructors employed in different universities and requested to share an online survey link with their students.

The survey was translated into Urdu – the national language of Pakistan. The English version was first translated into Urdu by a professional who had Master's degree in English. The translated Urdu version was then backwards translated into English by a second professional with similar qualifications. No discrepancies were found in the forward-backward translation of the survey, which enabled us to proceed with the data collection.

Table 3
Sample characteristics.

Items	Frequency	Percentage
Gender		
Male	424	41.2 %
Female	603	59.0 %
Age		
18-22	778	75 %
23-29	249	25 %
Study Level		
Undergraduate	766	74.6 %
Graduate	255	24.8 %

The survey consisted of four parts. The first part consisted of demographic questions followed by questions related to PMT constructs. The third part consisted of computer security behavior questions followed by smartphone security behavior questions. It should be noted that the survey questions were posed in such a way that both English and Urdu translations of each question were visible.

The research proposal for this study underwent scrutiny by the university's IRB. The minimum age of the participants who took part was 18 years of age, and it was ensured that they were not harmed physically, mentally, or psychologically while attempting the survey questionnaire. To ensure the anonymity and confidentiality of the participants, no personally identifiable information was collected. The students were given a consent form, which they undertook to voluntarily take part in the survey, and were given the option to withdraw any time they wanted.

A pilot study was conducted prior to the actual execution of the survey on a total of 50 participants. The first author of this study requested undergraduate students to take part in the pilot test. The students were contacted via email, and a lab was reserved for them, which had an Internet connection. The students were asked to fill out the questionnaire and ask any questions that they deemed incomprehensible. The first author of the study diligently observed the participants and noted down their queries. After the pilot test, the translation of a few of the questions was changed in such a way that it became easy for the students to understand.

The actual data were collected from April 2022 to December 2022. A total of seven universities took part in the survey. After the data collection, the responses were collected in the Excel sheet and underwent screening and cleaning, thereby reducing the sample size to 1027, which was subsequently used for final analysis. Table 1 presents the sample characteristics. As seen, a total of 59 % of participants were female with the majority of them (75 %) between the age group of 18-22 years old and studying at undergraduate level.

4.3. Data analysis

IBM SPSS V21 and covariance based – structural equation modeling (CB-SEM) in IBM AMOS version 21 was used for the explanatory modeling. CB-SEM is an appropriate technique when the aim of the research is theory testing (Kline, 2015). Further skewness and kurtosis was also performed on the data to check normality for ascertaining the appropriateness of SEM as analysis techniques. The skewness for SSB was -.049, indicating a nearly symmetric distribution. Similarly, the skewness for CSB was -.126, suggesting a slight negative skew but still approximating a normal distribution. These values are within acceptable ranges, and thus the data for both variables were considered to meet the assumption of normality. The distribution of both dependent variables (CSB and SSB) was assessed for normality using kurtosis. The kurtosis for Smartphone Security Behavior was 0.10, indicating a distribution that is nearly normal but with a slight tendency toward heavier tails and a sharper peak. The kurtosis for Computer Security Behavior was 0.018, suggesting a distribution that is essentially mesokurtic with a peak and tail shape close to that of a normal distribution. These values indicate that both variables approximate normal distributions. Therefore SEM was appropriate for carrying out data analysis. Structural Equation Modeling (SEM) is a robust data analysis technique utilized across various fields, offering benefits such as error control, mediation variable incorporation, and theoretical model evaluation.

CB-SEM is effective for testing theories, examining relationships between observed and latent variables (Anderson and Gerbing, 1988). We employed maximum likelihood estimation, which estimates parameters to produce a covariance matrix close to the observed covariance matrix. Goodness-of-fit evaluation is crucial, using indices like maximum likelihood estimation to assess model fit reliability.

Predictive modeling was carried out by using three machine learning (ML) algorithms in Python. Decision Trees (DT), K Nearest Neighbor (KNN) and Support Vector Machine (SVM). Decision Trees (DT) is a non-parametric supervised learning method used for classification and regression tasks. It creates a tree-like structure where each internal node represents a decision based on input features, and each leaf node represents an output value (Osisanwo et al., 2017). DTs are easy to interpret and handle categorical and numerical data well; hence, in our case, we used DT to identify the PMT constructs that significantly affect developing security intention. We also used KNN and SVM in our study. KNN is a simple and effective algorithm used for classification and regression tasks. It classifies data points based on the majority class of their nearest neighbors. KNN is easy to implement and can handle multi-class problems. Similarly, Support Vector Machine (SVM) constructs a hyperplane in a high-dimensional space to separate data points into different classes. SVM works well for both linear and non-linear data and is effective in high-dimensional spaces. SVM aims to maximize the margin between classes, which often leads to better generalization. However, SVM can be sensitive to the choice of the kernel and regularization parameters. In this case, 'poly' was used as a kernel along with the regularization parameter of 1 because it maps the data into a higher-dimensional space using polynomial functions and is effective for capturing nonlinear relationships. The advantages and disadvantages (Almazroi et al., 2020) of the three ML algorithms are given in Table 4.

We also investigated the significance of the PMT constructs using the wrapper feature selection approach to identify the top constructs that contributed the most towards correct prediction for computer and smartphone security behaviors (Kohavi and John, 1998) . The wrapper method iteratively trains a model on systematically chosen feature subsets and identifies the most significant features for accurate predictions. We also made use of principal component analysis (PCA) to facilitate the visualization of our data by mapping the multiple dimensions into two-dimensional space. PCA is a dimensionality reduction technique that projects high dimensional data into a lower dimensional space while retaining the most important patterns which simplifies the visualization and interpretation of complex datasets (Jolliffe, 2005).

5. Result and discussion

In this section, we present the results of the study. The explanatory modeling is reported in Section 5.1 whereas the results of predictive modeling are discussed in Section 5.2. The discussion is reported in Section 5.3 which includes the contribution of this study and implications for research.

5.1. Explanatory modeling

The explanatory modeling steps involve gauging the reliability and validity of the constructs which are given in Section 5.1.1. The measurement and structural model test were done by measuring the fit indices and is reported in Subsection 5.1.2. The third step is path analysis; which is explained in Subsection 5.1.3 for computer security and Subsection 5.1.4 for smartphone security.

Table 4
ML algorithm used in predictive analysis.

Algorithms	Advantages	Rationale for choosing
Decision Trees	The decision process is interpretable. The outcome can be traced back for better explainability	The visualization of the decision tree provides an easy explanation of the outcomes and can potentially reveal new insights into the data
K Nearest Neighbor	Best suited to numerical data Decisions are based on the proximity of the data points. Inherent support for non-linear decision boundaries	The proximity of data points is the basis for the initial clustering and labeling of the dataset hence KNN is expected to perform well on this data
Support Vector Machine	SVM provide better generalizability compared to other classifiers and more resilient to overfitting	SVM is known to work well in limited data scenarios while still providing strong accuracy

5.1.1. Reliability and validity

The factor loadings (Table 5) evaluate the correlation between the observed variables that define the same latent variable in order to determine whether or not the measurement model has convergent validity. In general, values of at least 0.3 and greater than 0.5 are considered to be satisfactory – (for computer security and smartphone measurement models are in Table 5) while values of more than 0.7 are considered to be very satisfactory (Hair et al., 2010). The measurement model's SMC (R^2) indicates the magnitude of the observed variable's variation that can be explained by the latent variables, according to (Sarstedt et al., 2016) items with a R^2 of less than 0.25 are likely to be removed, but in our case, there is no such item whose R^2 value is less than 0.25. The reliability and validity of the models were measured by Cronbach's alpha, average variance extracted (AVE) and composite reliability (CR), as shown in Table 6. The Cronbach's alpha or CR values greater than 0.7 are deemed appropriate, while AVE should be greater than 0.5 (Taber, 2018). As shown in Table 6, Cronbach's alpha and CR of the variables for both models (computer security and smartphone security) is greater than or equal to 0.7, which is consistent with the guidelines for instrument validation (Hair et al., 2010). The AVE values in this study are appropriate except for RC with AVE = 0.43 for computer security and AVE = 0.42 for the smartphone security model. The low AVE values show that almost 56 % of the variance for response cost is due to measurement error. However, the composite reliability being greater than 0.7 shows the reliability is intact (Fornell and Larcker, 1981).

5.1.2. Test of measurement and structural model

The measurement models for CSB and SSB were tested using fit statistics. This study reports absolute fit measures (AFM), parsimonious fit measures (PFM) and incremental fit measures (IFM) (Cheng, 2011; Hina et al., 2019), as shown in Table 7. The values of the goodness of fit index (GFI), comparative fit index (CFI), and tucker-lewis index (TLI) for smartphone security and computer security measurement models were greater than 0.9. The root mean square error (RMSEA) for both models is less than 0.08 which is also acceptable (Kline, 2015). There is no standard cutoff point for parsimony normed fit Index (PNFI) and parsimonious comparative fit index (PCFI) for determining a good fit, though (Cheng, 2011; Hina et al., 2019) state that an acceptable model is one with a value above 0.50. For this study, the PNFI was .798, and PCFI were .821, greater than the acceptable value of 0.5. The structural model test also shows a good fit. As shown in Table 7, the model fit indices for the computer security structural model were acceptable with [χ^2/df = 1778.45/474, GFI = .899, CFI = .916, TLI = .906]. The REMSEA for computer security is .052, and that of smartphone security is .051. The model fit indices for the smartphone security structural model were also acceptable with [χ^2/df = 1870.45/474, GFI = .904, CFI = .924, TLI = .914].

5.1.3. Explanatory analysis of computer security

As shown in Table 8, PMT was able to explain computer security behaviors The path coefficients show that in *threat appraisal* component of PMT, *threat severity* had a significant positive relationship with computer security intention ($\beta = 0.139$) with $p < .001$. This supports H2 (a). The *threat vulnerability* on the other hand did not show any significant relationship with computer security intention ($t = 1.766$, $p = 0.077$). In the *coping appraisal* component of PMT, *response efficacy* ($\beta = 0.221$), ($p < 0.001$) and *self-efficacy* ($\beta = 0.508$), ($p < 0.001$) were shown to have a significant positive relationship with computer security intention. Therefore, H3 (a) and H4 (a) are also supported. *Response cost*, on the other hand, had a negative relationship with computer security

Table 5
Variables, items and their loadings.

Computer Security	Measurement variables[1]	Item Loadings	Smartphone Security	Measurement variables	Item Loadings
Threat severity	PS1	0.74	*Threat severity*	PS1	0.70
	PS2	0.84		PS2	0.81
	PS3	0.83		PS3	0.84
	PS4	0.77		PS4	0.77
	PS5	0.69		PS5	0.70
	PS6	0.79		PS6	0.78
Threat Susceptibility	PV6	0.55	Threat Susceptibility	PV6	0.78
	PV5	0.77		PV5	0.79
	PV4	0.79		PV4	0.76
	PV3	0.77		PV3	0.67
	PV2	0.67		PV2	0.63
	PV1	0.62		PV1	1.17
Response cost	RC6	0.64	*Response cost*	RC6	0.62
	RC5	0.59		RC5	0.65
	RC4	0.66		RC4	0.66
	RC3	0.68		RC3	0.74
	RC2	0.72		RC2	0.63
	RC1	0.62		RC1	0.78
Response efficacy	RE4	0.66	*Response efficacy*	RE4	0.65
	RE3	0.79		RE3	0.79
	RE2	0.81		RE2	0.81
	RE1	0.61		RE1	0.62
Self-efficacy	SE6	0.76	*Self-efficacy*	SE6	0.70
	SE5	0.64		SE5	0.69
	SE4	0.69		SE4	0.70
	SE3	0.69		SE3	0.61
	SE2	0.64		SE2	0.53
	SE1	0.59		SE1	0.68
Security Intention	SI4	0.69	Security Intention	SI4	0.75
	SI3	0.76		SI3	0.73
	SI2	0.73		SI2	0.72
	SI1	0.72		SI1	0.63

[1] for item description, please consult Appendix.

Table 6
Reliability and validity of both models.

	Computer Security Model				Smartphone Security Model			
	Items	Cronbach's alpha	CR	AVE	Items	Cronbach's alpha	CR	AVE
TS	6	0.90	0.90	0.59	6	0.89	0.89	0.59
TV	6	0.90	0.85	0.50	6	0.85	0.85	0.50
RC	6	0.83	0.82	0.43	6	0.84	0.82	0.43
RE	4	0.80	0.81	0.52	4	0.81	0.82	0.52
SE	6	0.83	0.83	0.50	6	0.84	0.84	0.50
SI	4	0.81	0.81	0.52	4	0.82	0.82	0.51
Scale	16	0.83	0.81	-	14	0.87	0.86	-

Note: CR: Composite Reliability, AVE: Average Variance Extracted.

Table 7
Fit indices for measurement and structural model.

	AFM				IFM		PFM	
	χ^2/df	GFI	AGFI	RMSEA	CFI	TLI	PNFI	PCFI
Measurement Models								
CSB	1648.097/442	0.90	0.88	0.05	0.92	0.91	0.80	0.82
SSB	1606.978/442	0.90	0.89	0.05	0.92	0.92	0.80	0.82
Structural Models								
CSB	1778.454/474	0.90	0.88	0.05	0.92	0.91	0.80	0.82
SSB	1870.950/474	0.90	0.89	0.05	0.92	0.91	0.80	0.82

Note: AFM: absolute fit measures, IFM: incremental fit measures, PFM: parsimonious fit measures, PFI: parsimonious fit index, GFI: goodness of fit index, RMSEA: root mean square error, CFI: comparative fit index, TLI: tucker-lewis index, PNFI: parsimony normed fit Index, PCFI: parsimonious comparative fit index, CSB: computer security behavior, SSB: smartphone security behavior.

intention which was insignificant ($p = 0.137$). However, the low AVE value of RC should be noted, with almost 57 % of the variance being attributed to measurement error. The coefficient of determination R^2 for computer security intention is 0.73, and it shows that 73 % of the variation can be explained by *threat severity, response efficacy* and *self-efficacy*. The intention to secure computer device also had a strong significant relationship with computer security behavior with ($\beta = 0.590$), ($p < 0.001$) supporting H6 (a). The R^2 for computer security behavior is 0.168, depicting 16 % of the variance in computer security behavior, which is to be explained by computer security intention. The full

Table 8
Hypothesis testing of computer security behavior.

Hypothesis	Path	Path Coefficients (β)	t-value	p-value	Supported
H1(a)	*Threat vulnerability* -> Intention to Secure Device	0.05	1.77	0.08	No
H2(a)	*Threat severity* -> Intention to Secure Device	0.14	5.01	<0.001	Yes
H3(a)	*Response efficacy* -> Intention to Secure Device	0.22	5.76	<0.001	Yes
H4(a)	*Self-efficacy* -> Intention to Secure Device	0.51	14.89	<0.001	Yes
H5(a)	*Response cost* -> Intention to Secure Device	-0.04	-1.49	0.14	No
H6(a)	Intention to Secure Device -> Computer Security Behavior	0.59	12.03	<0.001	Yes

regression model is presented below;

Intention to Secure Device = 0.05 X *Threat vulnerability* + 0.14 X *Threat severity* + 0.51 X *Self-efficacy* + 0.22 X *Response efficacy* – 0.04 X *Response cost*

5.1.4. Explanatory analysis of smartphone security

As shown in Table 9, PMT was able to explain smartphone security behaviors. In the same vein as the computer security PMT model, the path coefficients for smartphone security show that the *threat appraisal* component of PMT *threat severity* ad a significant positive relationship with smartphone security intention ($\beta = 0.102$) with $p < 0.001$. Therefore, H2 (b) is supported. The *threat vulnerability*, on the other hand, did not show any significant relationship with smartphone security intention ($t = 1.837$, $p = 0.066$). In the *coping appraisal* component of PMT, *response efficacy* ($\beta = 0.196$), ($p < 0.001$) and *self-efficacy* ($\beta = 0.562$), ($p < 0.001$) were shown to have a significant positive relationship with smartphone security intention. Therefore, H3 (b) and H4 (b) are supported. Response cost, on the other hand, had a negative relationship with smartphone security intention, which was insignificant ($p = 0.225$). Here, too, the low AVE value of RC should be noted, with almost 57 % of the variance being attributed towards measurement error. The coefficient of determination R^2 for smartphone security intention is 0.89, showing that 89 % of the variation can be explained by threat severity, response efficacy and self-efficacy. The intention to secure smartphone device also had a strong significant relationship with smartphone security behavior with ($\beta = 0.528$), ($p < 0.001$), thereby H6 (b) is also supported. The R^2 for smartphone security behavior is 0.13, depicting 13 % of the variance in smartphone security behavior to be explained by smartphone security intention. The full regression model is presented below;

Intention to Secure Device = 0.06 X *Threat vulnerability* + 0.10 X *Threat severity* + 0.56 X *Self-efficacy* + 0.20 X *Response efficacy* – 0.03 X *Response cost*

Table 9
Hypothesis testing of smartphone security behavior.

Hypothesis	Path	Path Coefficients (β)	t-value	p-value	Supported
H1(b)	*Threat vulnerability* -> Intention to Secure Device	0.06	1.84	0.07	No
H2(b)	*Threat severity* -> Intention to Secure Device	0.10	3.78	<0.001	Yes
H3(b)	*Response efficacy* -> Intention to Secure Device	0.20	5.19	<0.001	Yes
H4(b)	*Self-efficacy* -> Intention to Secure Device	0.56	15.13	<0.001	Yes
H5(b)	*Response cost* -> Intention to Secure Device	-0.03	-1.22	0.22	No
H6(b)	Intention to Secure Device -> Smartphone Security Behavior	0.53	10.38	<0.001	Yes

5.2. Predictive modeling

The predictive modeling was done at two levels – macro and micro. The macro level predictive modeling was done to understand the predictive power of PMT for complete computer security and smartphone security behaviors. The micro level predictive modeling was done to understand the predictive power of PMT for the dimensions of computer and smartphone security behaviors. To get a more nuanced understanding, the micro level involved the prediction of four types of computer behavior, i.e. device security, updating behavior, proactive awareness and password protection, and two types of smartphone security behavior, i.e. technical Security and social security. Clustering each dimension separately allows us to explore specific patterns within individual aspects, highlighting areas where security practices differ significantly or remain consistent across the dataset. At the macro level, the prediction of computer security behavior was made by making use of the Computer Security instrument - SeBISs'16 items (Subsection 5.2.1) - and smartphone security instrument SSBSs' 14 items (Subsection 5.2.2). Our dual macro and micro predictive modeling enables us to both drill down into specific security practices and understand the broader security landscape, ensuring that our analysis is comprehensive and actionable.

5.2.1. Computer security predictive modeling

For macro level predictive analysis, we pre-processed the data to discretize the computer security behavior. This was done by finding clusters, as the literature on cybersecurity does not provide any guidelines for the classification. This led us to find a natural grouping in the data based on the security behavior of the respondents. We employed the K-means clustering algorithm to discover the hidden clusters in our data as a basis for class labeling. Since the number of these groups was not known initially, the elbow method yielded two as the optimal number of clusters (depicted in Fig. 3). We utilized the elbow method in our study because it is a well-established and widely accepted strategy for determining the optimal number of clusters (Hamka and Ramdhoni, 2022). The elbow method balances model complexity against explained variance and provides a clear visual indication of where adding more clusters yields diminishing returns. This approach ensures that we capture the most meaningful structure in the data while avoiding overfitting. Although the Pareto principle could simplify the model by focusing on clusters that explain the majority of variance, the elbow method provides a more data-driven approach, aligning well with our goal of identifying natural groupings based on the overall data distribution. The initial clustering was based on all 16 items embodying different aspects of security behavior taken as features. For the sake of visualization, the 16-dimensional data was transformed into two-dimensions using principal component analysis and is shown in Fig. 3. The two clusters exist in close proximity with some overlapping at the inner boundary but greater scatter at the farther ends. The dataset

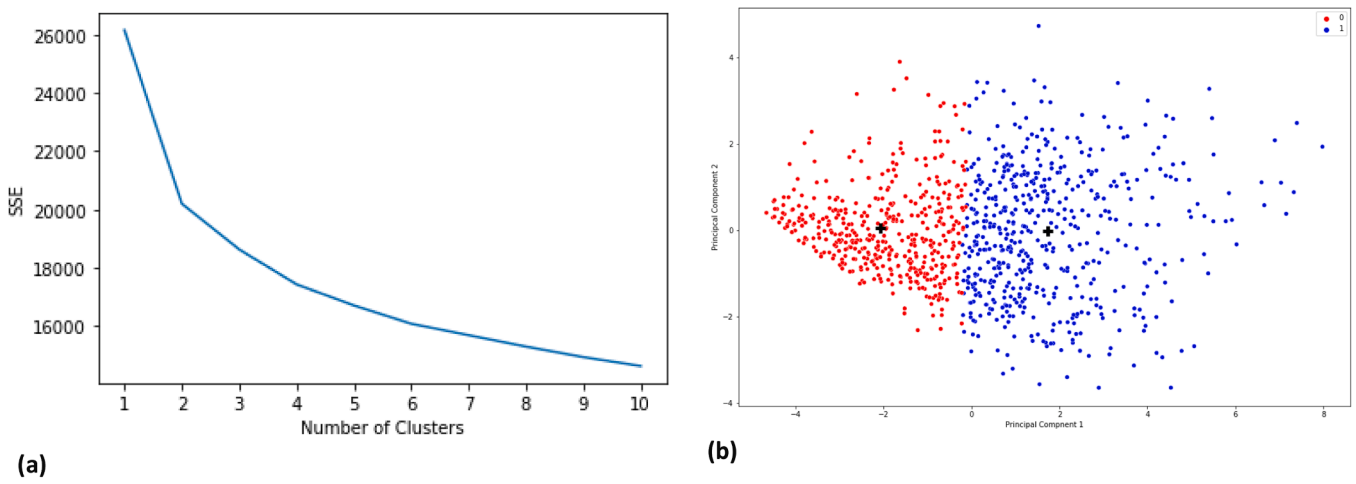

Fig. 3. Computer security behavior (a) elbow method, (b) clusters.

was labeled according to the cluster membership as 0 for Red and 1 for blue (Table 10). After clustering, the protection motivation theory's components were employed to predict the cybersecurity practices. We argue that if the components of PMT can predict the correct class label for instance, then it confirms the relationship between PMT and computer security behavior, as the theory predicts.

Three machine learning algorithms were used to see the predictive accuracy of protection motivation in predicting the correct class labels for the behavior of an individual. The three classifiers – decision tree (DT), K nearest neighbor (KNN) and support vector machine (SVM) – predicted the computer security behavior and an accuracy measure was used to gauge the predictability power. As can be seen from Table 11, the highest accuracy of 76 % was achieved by KNN followed by SVM (71 %) and DT (65 %). The average accuracy from the three classifiers remained 70 % which depicts the significant predictive power of PMT in identifying the correct class label of computer security behavior. To see detailed prediction and accuracy of classification, confusion matrices are used which are shown in Fig. 4. The confusion matrix is the summarized detail of prediction results showing the total counts of a specific class (Ling et al., 2003).

We extended our analysis to understand the predictability for specific types of computer security behavior at a micro level. As has been discussed there are four dimensions of computer security instrument SeBIS namely Device security, Upgrading, Proactive Awareness, and Password Generation – each of which was taken as a specific computer security behavior in micro level predictive analysis. Therefore, the dataset was broken down into four datasets, with each data set only containing computer security items relevant to the specific security behavior which is also called dimension. For instance, the dataset relevant to the device security dimension contained only 4 items of SeBIS, whereas that of proactive awareness dimension contained 6 items. As a first step, k- means clustering algorithm was employed to find clusters for each type of computer security behavior. It should be noted that the clustering algorithm was run 4 times on the respective (DS, UP, PA, PG) dimensions' data set. The results of dimension-based clustering are shown in Fig. 5 which exhibit noticeable differences in the grouping and patterns. Clusters based on devise security behavior (Fig. 5 a) reveal the presence of micro clusters fairly evenly distributed across the range. This is an indicator of unique security personas relative to device security behavior. The personas appear even more distinctive in case of updating behavior-based clustering (Fig. 5 d) but relatively cluttered with password generation-based behavior (Fig. 5 b). The proactive awareness behavior (Fig. 5 c) yields clusters very similar to those with the complete computer security feature set including all behaviors. This suggests higher weightage of this type of behavior (PA) by the clustering algorithm. Table 10 presents numerical summaries of the clusters formed for each dimension of computer security behavior as well as the complete SeBIS dataset. We also show the probability distribution graphs of the clusters for a quick visual comparison of the central tendency, spread, and likelihood of different scores between the two clusters. The probability density graph (PDG) of the mean scores of the clusters (Figs. 6 d, 7 d, 8 d, 9 d) show significant overlapping in all the clusters except the one with complete SeBIS (Fig. 4 d). This is a clear indication of the fact that the clustering algorithm segregated the instances based on similarities in the behavior and not the aggregate score of an individual.

After the careful clustering of the four datasets, three machine learning algorithms were used to see the predictive accuracy of protection motivation in forecasting the correct class labels for four types of computer security behavior of an individual. It should be noted that the algorithm was run four times on each data set (each dataset of a specific dimension of SeBIS). The highest accuracy (78.38 %) was observed for device security (DS) behavior by KNN classifier followed PG dataset at 75.89 %, as shown in Table 11. KNN consistently performed better for all five datasets of computer security behavior compared to other classifiers. From the dataset perspective, the highest average accuracy was

Table 10
Statistics for clusters formed for each computer security behavior.

Dataset	Numerical statistic	Red cluster	Blue cluster
Complete SeBIS	Mean Aggregate	3.94	2.87
	Std Aggregate	0.38	0.44
	Max Aggregate	4.71	3.47
	Min Aggregate	3.24	1.00
	Count	466	561
Device Security dimension	Mean Aggregate	3.62	2.88
	Std Aggregate	0.59	0.55
	Max Aggregate	4.71	4.47
	Min Aggregate	1.94	1.00
	Count	646	381
Password Generation dimension	Mean Aggregate	3.81	2.97
	Std Aggregate	0.51	0.55
	Max Aggregate	4.71	4.29
	Min Aggregate	2.29	1.00
	Count	469	558
Proactive Awareness dimension	Mean Aggregate	3.80	2.89
	Std Aggregate	0.49	0.50
	Max Aggregate	4.71	4.17
	Min Aggregate	2.29	1.00
	Count	521	506
Updating dimension	Mean Aggregate	2.91	3.77
	Std Aggregate	0.56	0.56
	Max Aggregate	4.41	4.76
	Min Aggregate	0.94	2.17
	Count	494	533

Table 11
Prediction accuracy of computer security behaviors.

	DT		KNN		SVM		Mean
Dataset	Accuracy (%)	Features	Accuracy (%)	Features	Accuracy (%)	Features	
Complete SeBIS	65.78	['TV' 'RC' 'SI']	76.00	['RC' 'SE' 'SI']	71.11	['RC' 'SE' 'SI']	70.96
DS	61.00	['RE' 'SE' 'SI']	78.38	['TS' 'RE' 'SI']	71.04	['TS' 'RE' 'SI']	70.14
PG	63.84	['TS' 'RC' 'SE']	75.89	['TV' 'RC' 'SI']	69.64	['TV 'RC' 'SI']	69.79
PA	59.81	['TS' 'RE' 'SE']	69.86	['TS' 'TV' 'SE']	65.07	['TS' 'RC' 'SE']	64.91
UP	59.81	['TS' 'SE' 'SI']	73.36	['TS' 'SE' 'SI']	63.08	['RE' 'SE' 'SI']	65.42
Mean	62.05		74.70		67.99		68.24

NOTE: DS: Device Security, PG: Password Generation, PA: Proactive Awareness, UP: Updating, TV: Threat vulnerability, TS: Threat severity, RE: response efficacy, RC: Response cost, SI: Intention to secure devic, DT: Decision Tree, KNN: K nearest neighbor, SVM: Support vector machine.

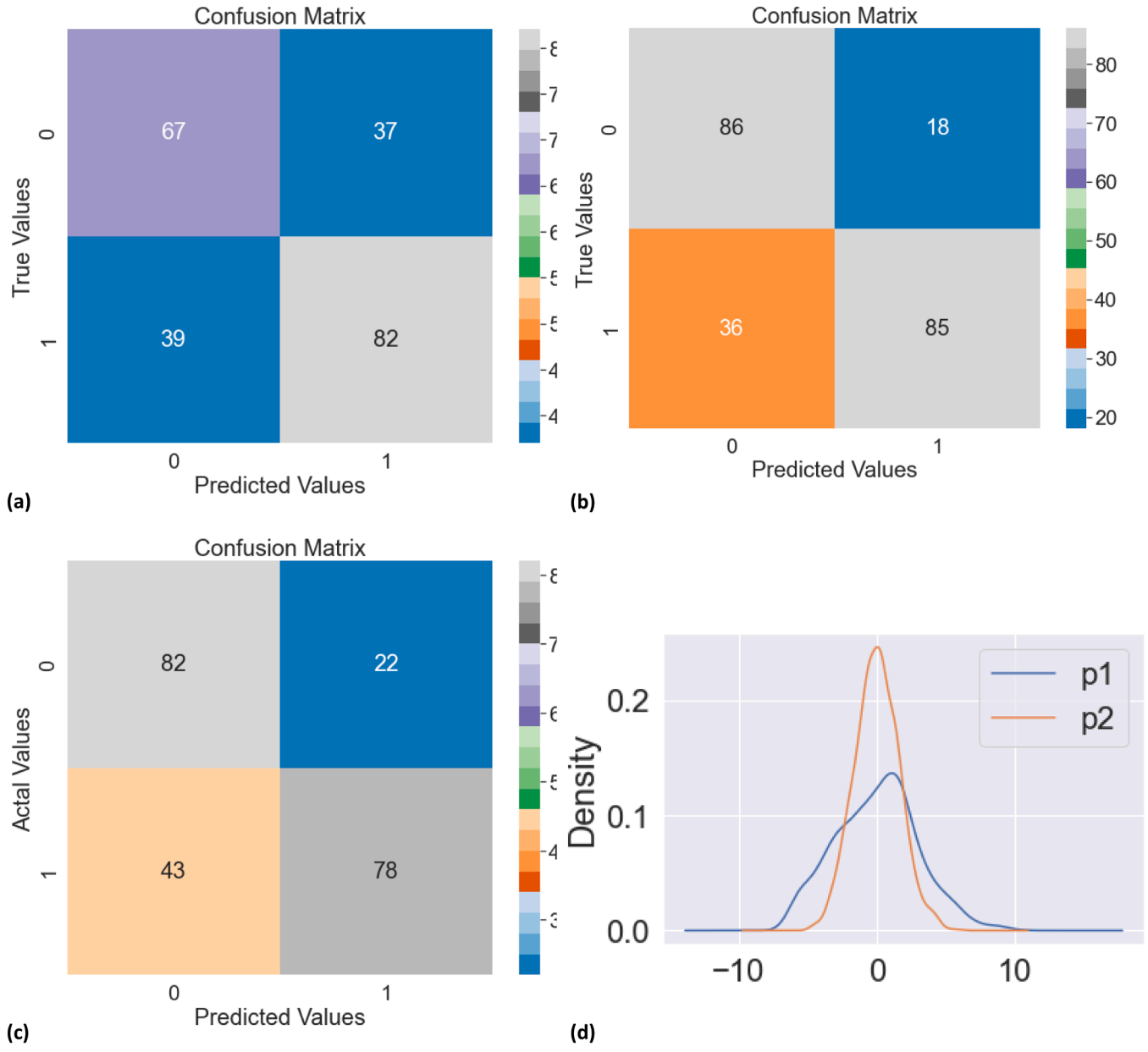

Fig. 4. Confusion matrix for complete computer security behavior for (a) DT, (b) KNN, (c) SVM; (d) PDG of principal components.

70.96 for complete SeBIS dataset, followed by 71.14 % for DS dimension. The average prediction accuracy was also calculated for the three ML algorithms. The highest average accuracy was that of KNN (74.38 %), followed by SVM (67 %) and DT (62 %). Moreover, average accuracy was also computed (by taking means of DT, KNN and SVMs predictive accuracy) for findings on the predictive power of PMT in predicting computer security behaviors. As seen in Table 11, the average predictive accuracy of PMT for complete computer security behavior is 70.96 %. The overall predictive accuracy of PMT for device security is 70.14 % while for updating behavior is 65.42 %. The average accuracy from all

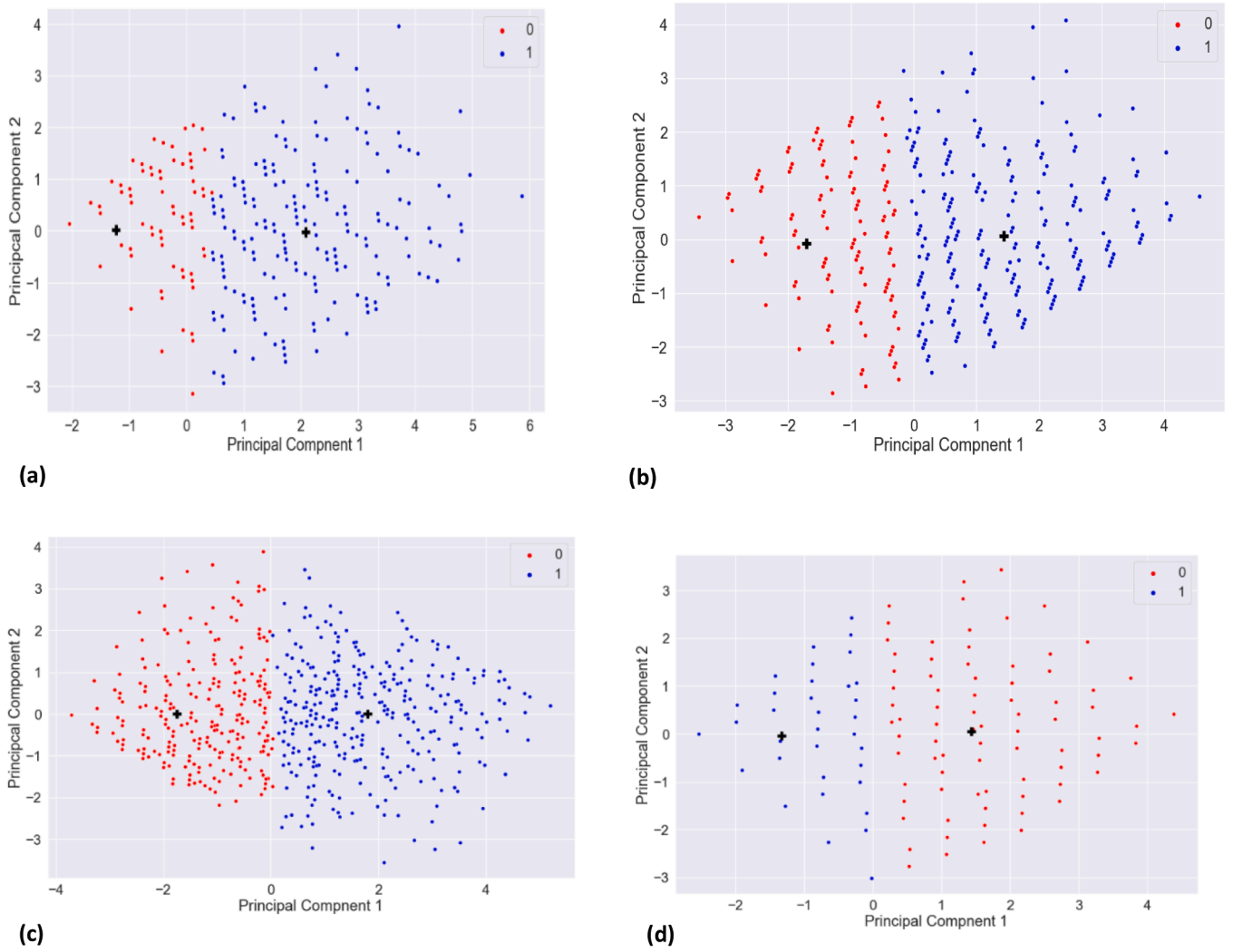

Fig. 5. Clusters for (a) device security, (b) password generation, (c) proactive awareness, (d) updating behaviors.

classifiers against all datasets remained above 60 % which depicts substantial predictive power of PMTs components in identifying the correct class label for different types of computer security behavior. That means, since the average accuracy was significantly higher than the baseline accuracy of 50 % (chance outcome for binary classification), it signifies the reliability and stability of the features to classify the computer security model in an independent manner. The confusion matrices for DS dataset at shown in Fig. 6; for PG dataset in Fig. 7; for PA dataset in Fig. 8 and for UP dataset in Fig. 9.

We also investigated the feature significance using the wrapper feature selection approach to identify top 3 PMT features that contributed significantly towards the correct prediction from the classifier in computer security behavior. This was done to study how well the predictive models aligned with SEM explanatory modeling results. The selected features are shown in Table 11 against each dataset of computer security behavior for respective classifiers. The *self-efficacy* and security intention features were selected most frequently by the classifiers as significant, however, the overall ranking of features showed significant variation between different datasets and classifier combinations.

5.2.2. *Smartphone security predictive modeling*

For smartphone security predictive modeling, we employed the same approach as computer security predictive modeling i.e. macro level and micro level ML analysis. At macro level, pre-processing of data was done to discretize the complete smartphone security behavior (containing all 14 items). Elbow method was used to find the number of clusters (Fig. 10 a) that were naturally occurring in the data. K-mean clustering algorithm applied on all 14 items yielded 2 clusters (Fig. 10 b) and the dataset was labeled as per the cluster membership 0 for red and 1 for blue. Three machine learning algorithms (DT, SVM and KNN) were then employed to find the predictive accuracy of PMT in forecasting correct class labels for smartphone security behavior (Table 12). The highest accuracy was achieved by KNN (68 %) followed by SVM (63 %). The average accuracy for the three classifiers was 62.14 %. The confusion matrices for complete smartphone security behavior are given in Fig. 11.

At micro level, each dimension of smartphone security behavior was taken as a specific smartphone security behavior. There are two dimensions of smartphone security behavior scale namely *Technical Dimension* and *Social Dimension*. The dataset was broken down into two datasets containing relevant items for technical and social dimensions of smartphone security behavior. Each dataset underwent K-mean clustering algorithm to ascertain clusters for each type of smartphone security behavior. The results of dimension based smartphone security clustering are shown in Fig. 12. After clustering of the two datasets, machine learning algorithms were employed for the predictive accuracy of PMT in predicting correct class labels. The highest accuracy of 75 % was observed for the social dimension of the smartphone security behavior by KNN classifier, and it consistently performed better for all three datasets (complete smartphone security, T and S dimension of smartphone security) in smartphone security prediction. The average

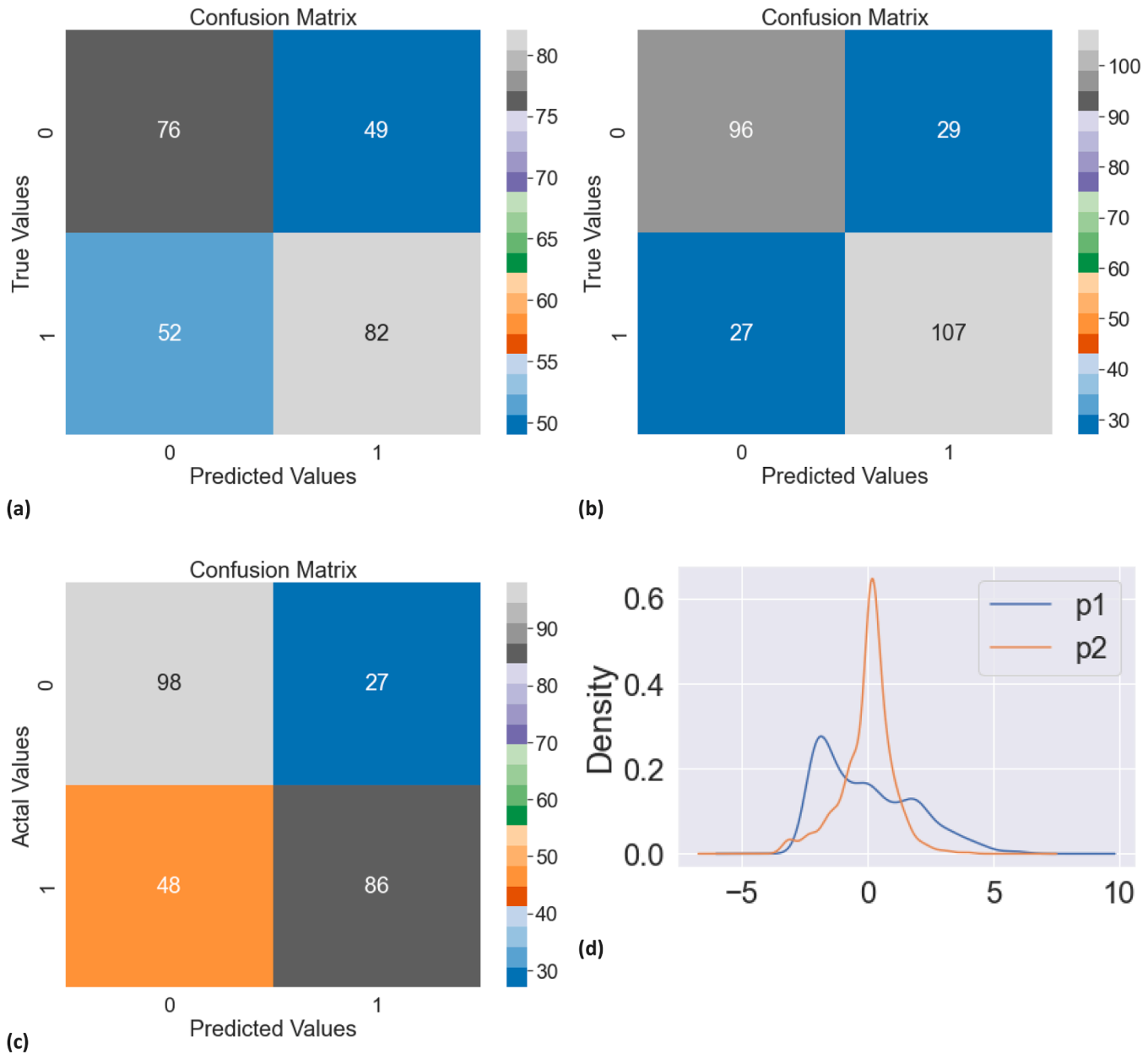

Fig. 6. Confusion matrices for device security (a) DT, (b) KNN, (c) SVM; (d) PDG principal components of device security.

accuracy for the ML algorithm was calculated, and again, the highest was that of KNN (71.85 %). The average accuracy of the three algorithms for predicting smartphone security behavior was 62 %. Whereas the average accuracy values for the technical smartphone security dimension and social smartphone security dimension were 67 % and 70 %, respectively. Again, since the average accuracy was significantly higher than the baseline accuracy of 50 % (chance outcome for binary classification), it signifies the reliability and stability of the features to classify the smartphone security models in an independent manner. Nevertheless the lowest average accuracy against the three datasets was at least 60 %, ruling out the case of chance random predictions. The confusion matrices for *the Technical Dimension dataset are given in Fig. 13, while those of the* S dataset are given in Fig. 14. The probability density graphs (PDG) of the mean scores of the clusters for smartphone security behavior are shown in Figs 11 d, 13 d and 14 d.

We also investigated the feature significance using the wrapper feature selection approach to identify the top three PMT features that contributed significantly towards the correct prediction from the classifier in smartphone security. This was done to study how well the predictive models aligned with SEM explanatory modeling results of smartphone security PMT models. The selected features are shown in Table 12 against each dataset for respective classifiers. The SE and SI features were selected most frequently by the classifiers as significant in smartphone security behavior, however, the overall ranking of features showed slight variation between different datasets (complete, T and S smartphone security behaviors) and classifier combinations.

NOTE: DS: Device Security, PG: Password Generation, PA: Proactive Awareness, UP: Updating, TV: Threat vulnerability, TS: Threat severity, RE: response efficacy, RC: Response cost, SI: Intention to secure devic, DT: Decision Tree, KNN: K nearest neighbor, SVM: Support vector machine.

6. Discussion

In the current study, predictive modeling is employed following a data-driven approach using ML algorithms to augment the results of explanatory modeling of PMT. This is distinctive as it reflects the

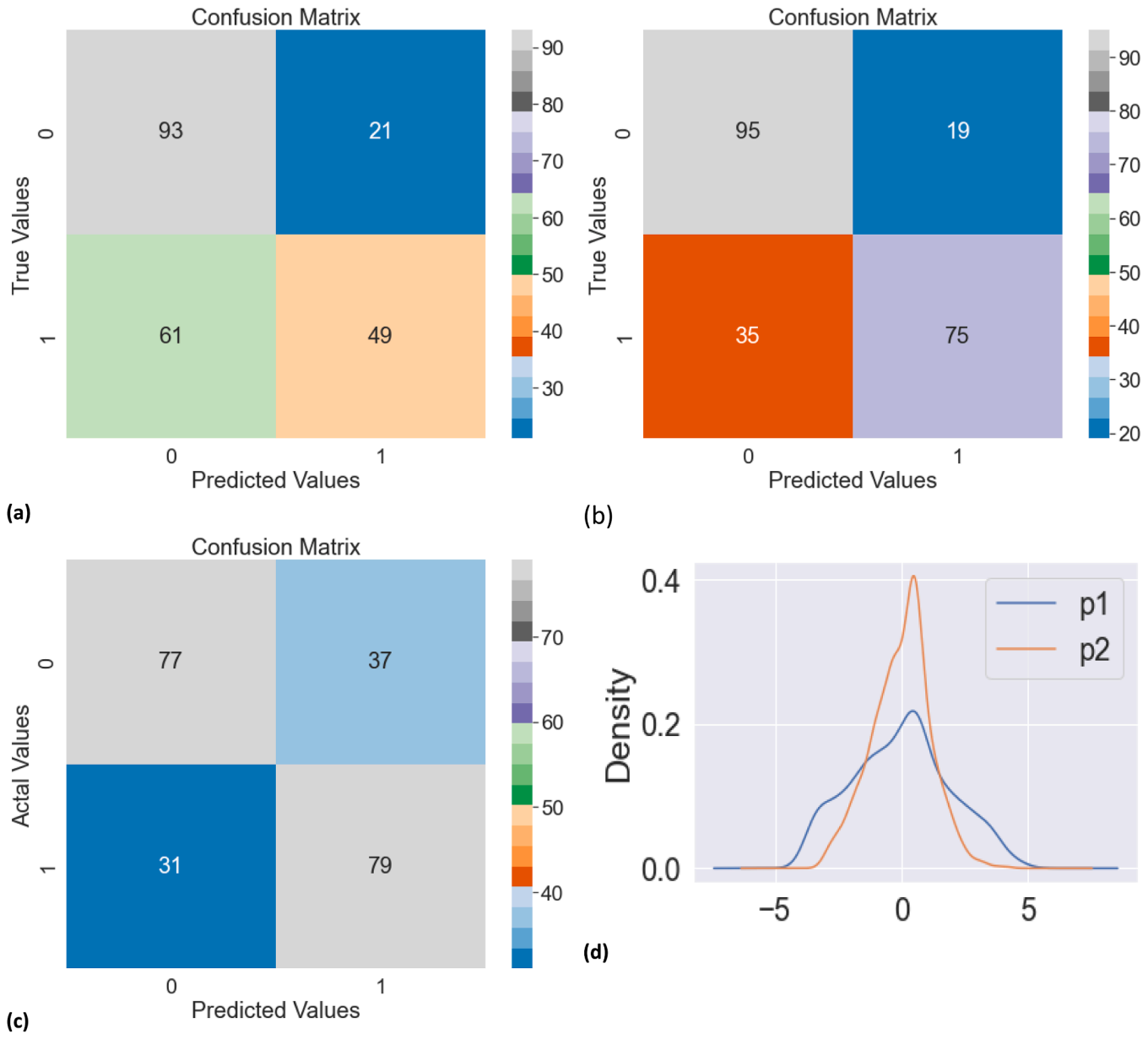

Fig. 7. Confusion matrices for password generation (a) DT, (b) KNN, (c) SVM; (d) PDG principal components of password generation.

underlying data without taking into consideration the assumptions in explanatory modeling. To ascertain the optimal performance of the PMT models for computer and smartphone security, a combination of linear and nonlinear ML algorithms are utilized. The combination of SEM and prediction analyses bridges the theory-practice gap and answers the research call of (Alassaf and Alkhalifah, 2021).

6.1. *Overview of the findings*

The explanatory modeling results of our study related to TV are in contrast to (Dang-Pham and Pittayachawan, 2015; Haag et al., 2021; Posey et al., 2015; Rajab and Eydgahi, 2019; Thompson et al., 2017), whereas the results of TS in our study are in line with (Hovav and Putri, 2016; Knapova et al., 2021; Thompson et al., 2017; Verkijika, 2018). The results related to *Response efficacy* in our study are in line with (Haag et al., 2021; Ifinedo, 2012; Rajab and Eydgahi, 2019; Vance et al., 2012; Vrhovec and Mihelič, 2021) (Dang-Pham and Pittayachawan, 2015; Giwah et al., 2019) while that of *self-efficacy* of our study are in line with (Dang-Pham and Pittayachawan, 2015; Giwah et al., 2019; Knapova et al., 2021; L. Li et al., 2019; Posey et al., 2015; Thompson et al., 2017; Verkijika, 2018). The results of *response cost* in our study are in contrast to that of (Dang-Pham and Pittayachawan, 2015; Hovav and Putri, 2016; Posey et al., 2015; Thompson et al., 2017; Tsai et al., 2016; Vance et al., 2012; Verkijika, 2018) and may be attributed towards the low AVE value of the response cost in our models. The explanatory power of PMT for computer security intention is 73 % and for security behavior is 16 % while for smartphone security intention is 89 % and for smartphone security behavior is 13 %, as ascertained by R^2.

The predictive modeling results reveal that the predictive accuracy for computer security behavior is 70 % as we averaged out the predictive accuracy for the three algorithms employed. For different types of computer security behavior, the accuracy is approximately between the range of 65 % - 70 % for device security, proactive awareness, password generation and updating behavior. This shows that there is little variation in predicting different types of computer security behaviors using PMT. On the other hand, the predictive accuracy is 62 % for smartphone security. However, the accuracy for different types of smartphone security behaviors is higher with 67 % for technical security for

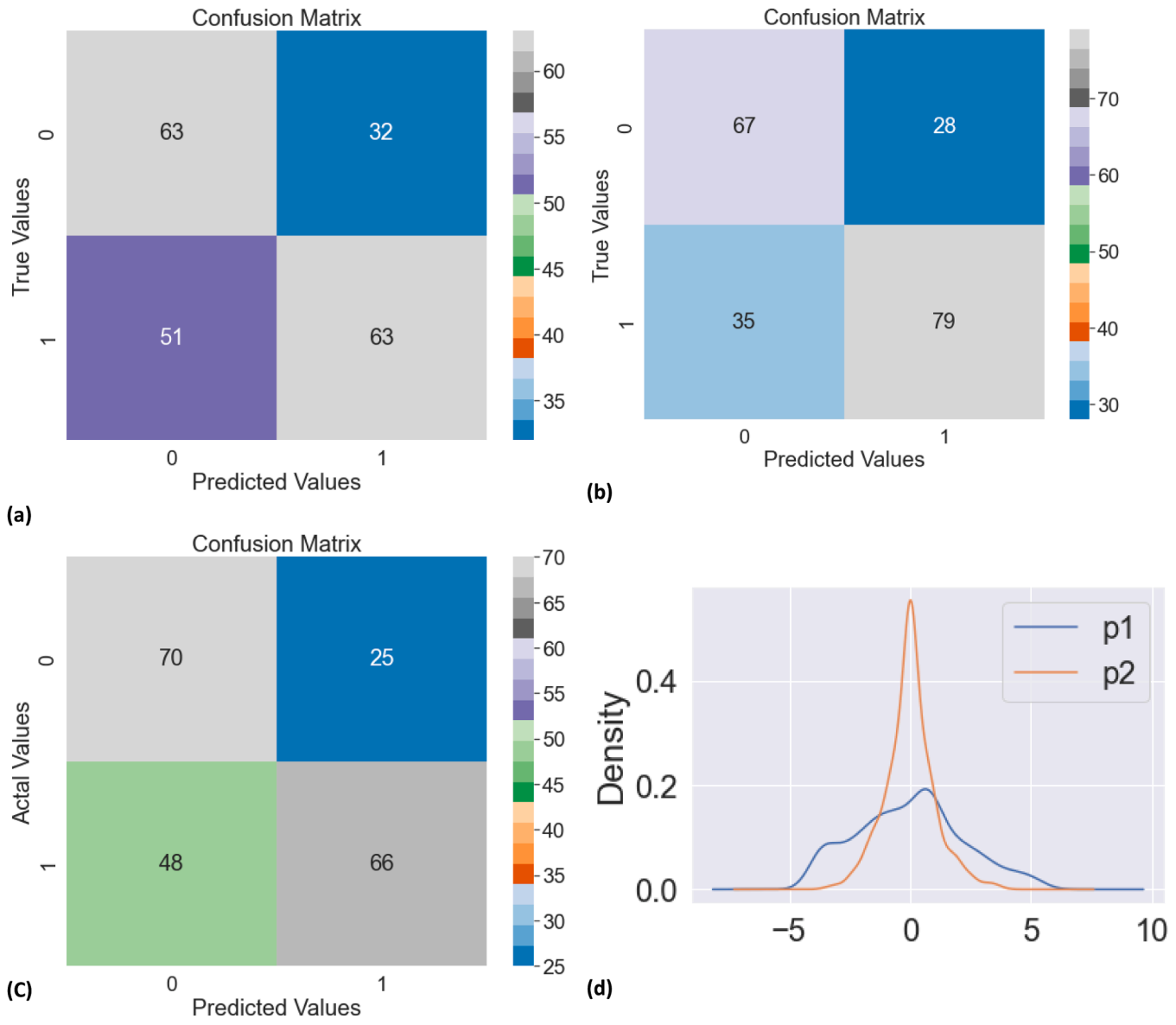

Fig. 8. Confusion matrices for proactive awareness (a) DT, (b) KNN, (c) SVM; (d) PDG principal components of proactive awareness.

smartphone and 70 % for social security for smartphone. The best performing algorithm for both computer and smartphone security is KNN, however the predictive accuracy does not vary drastically across different ML algorithms. These results also highlight that the interaction between PMT elements are not linear and PMT model is best estimated by the application of linear as well as nonlinear algorithms. The predictive power of PMT models for cyber security behaviors reveal that our understanding of the phenomenon under the lens of PMT is limited and other factors such as fear appeals, need to be incorporated.

These limitations of the PMT explanatory modeling is shown in the results from the wrapper method. On one hand it was consistently shown that *perceived threat severity, self-efficacy* and *response efficacy* were to be the most important features in prediction of security behaviors. These findings echo the observations made by the previous studies that the *coping appraisal* dimension is the most significant predictor of computer security behavior when compared to the *threat appraisal*. However, the inclusion of *response cost* and perceived vulnerability by some ML algorithms for various cybersecurity behaviors highlight the overall importance of PMT constructs. As observed in previous literature, the self confidence in the effective usage of security controls minimizes the perceived severity of the threats. According to the previous literature (Boss et al., 2015; Posey et al., 2015), the *coping appraisal* process of the protection motivation theory works when the *self-efficacy* and *response efficacy* are higher as compared to the *response cost* (RC) for the individuals to have intention to secure themselves. This study echoes this observation that *response efficacy* (RE) and *self-efficacy* may render the RC associated with protecting the devices insignificant. Nevertheless, it should also be noted that the AVE for *response efficacy* was below the accepted value of 0.5 that could potentially introduce measurement errors. Moreover, the study population being university students also play role in insignificant influence of response cost due to their reckless attitude (Arnett, 1996). Our results are also corroborated by a recent meta-analytical based theory testing of PMT (Mou et al., 2022) in which the coping-appraisal component had stronger influence than the threat-appraisal.

The *response cost* and perceived vulnerability to be the antecedents by some ML algorithms shows the importance of all the PMT constructs. This shows that the antecedents in PMT models are not compensatory in nature rather the theory is sound for its coping as well as *threat appraisal*. As shown by our analysis the non-existence influence of *response cost* and perceived vulnerability is not compensated by the significant influence of *threat severity, self-efficacy, response efficacy* and security intentions in

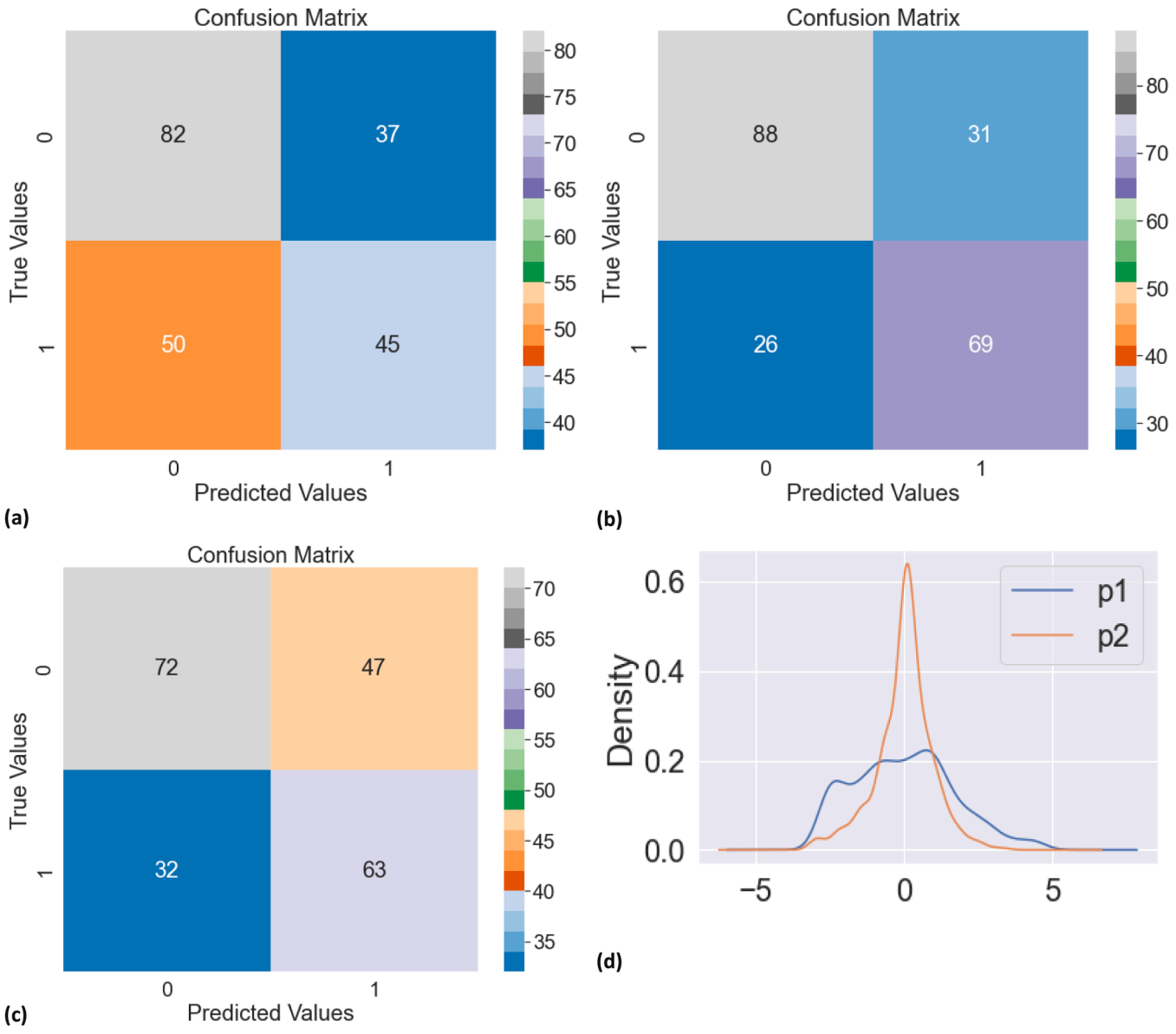

Fig. 9. Confusion matrices for updating (a) DT, (b) KNN, (c) SVM; (d) PDG principal components of updating.

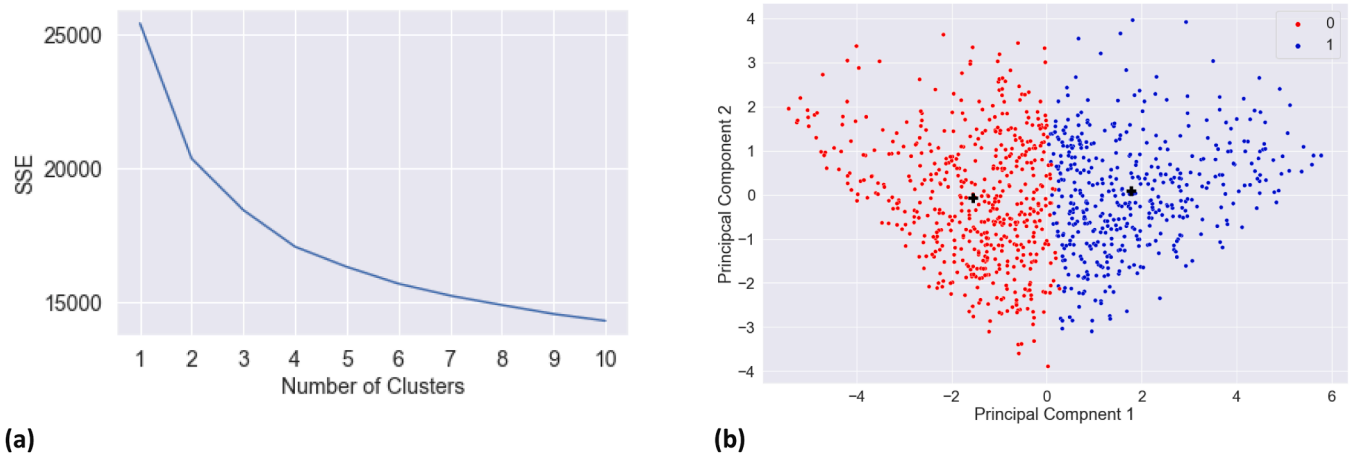

Fig. 10. Smartphone security behavior (a) elbow method, (b) Clusters.

Table 12
Prediction accuracy of smartphone security behavior.

	DT		KNN		SVM		Mean
Dataset	Accuracy (%)	Features	Accuracy (%)	Features	Accuracy (%)	Features	
Complete	54.37	['RC' 'SE' 'SI']	68.45	['TV' 'RC' 'SI']	63.59	['TS' 'RE' 'SI']	62.14
Technical Dimension	61.68	['RC' 'SE' 'SI']	71.84	['TS' 'TV' 'SE']	68.47	['RC' 'SE' 'SI']	67.33
Social Dimension	67.96	['TV' 'RE' 'SI']	75.27	['TS' 'RE' 'SI']	69.42	['TS' 'RE' 'SI']	70.88
Mean	61.34	-	71.85	-	67.16	-	66.78

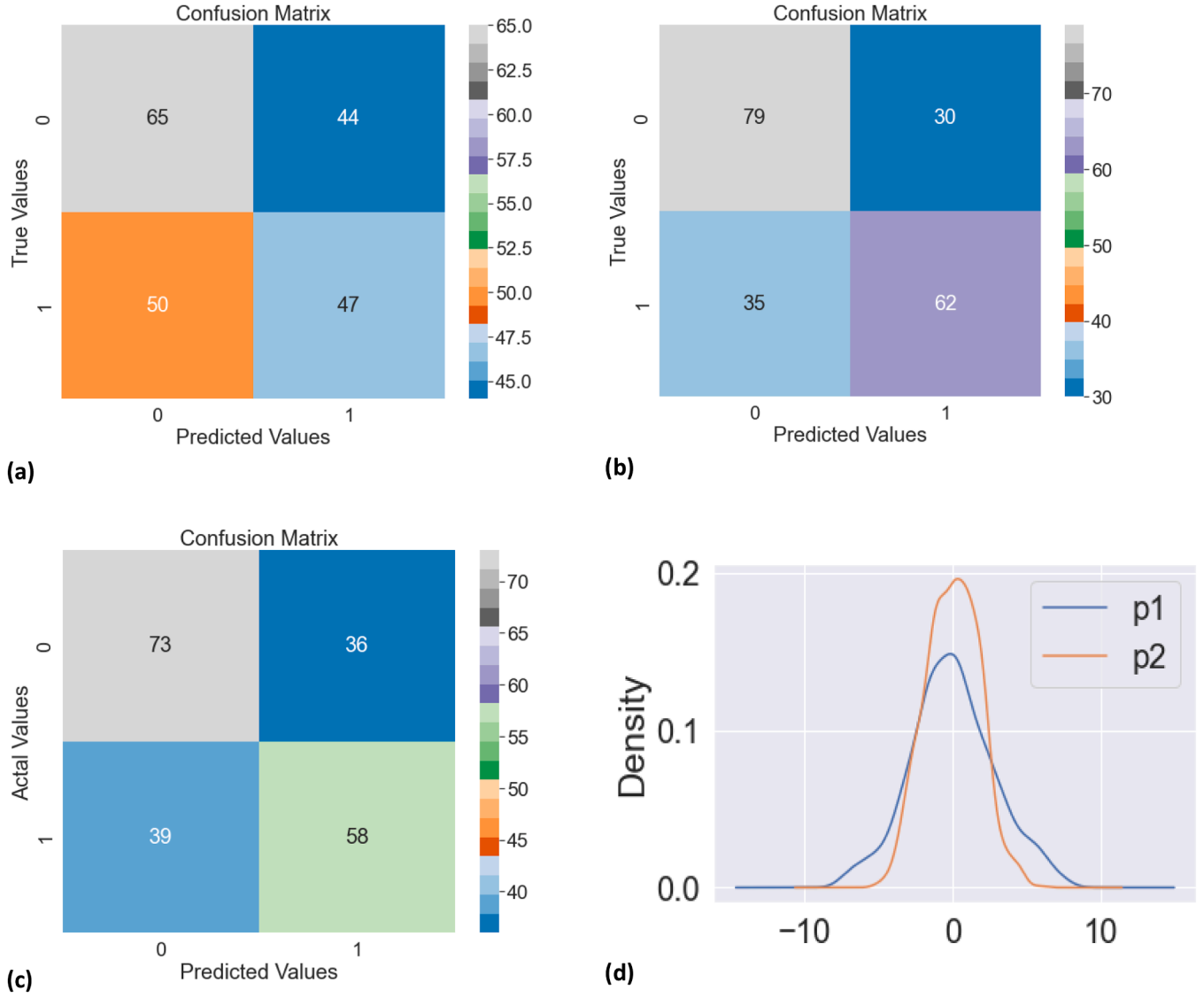

Fig. 11. Confusion matrices for smartphone security behavior (a) DT, (b) KNN, (c) SVM; (d) PDG principal components of smartphone security behavior.

both computer and smartphone security models. Nevertheless, the features that were the most important in PMT models were *self-efficacy*, security intention and *response efficacy*. These were the antecedents that were considered most important by that ML algorithms more frequently than the others. This as a result validates the protection motivation theory in explaining and predicting cybersecurity behaviors. One thing to note here is that, the 70 % predictive accuracy of PMT shows that there is still a need for carrying out more research by adding additional factors to the PMT model as has been argued by (Shmueli and Koppius, 2011).

6.2. *Synthesis of the two methods*

This study reveals the importance of two methods in understanding cyber security behavior phenomenon. The explanatory modeling is limited to overfitting of data due to fitting the data to explain past events, incorrect inferences due to correlational assumptions between variables, parsimony principles upholding and inability to predict future events and hence are not suitable for predictive validation of theories (Shmueli and Koppius, 2011). On the other hand, predictive modeling is limited in terms of offering causal insights and relationship between the variables causing inability for supporting or refuting claims about causality in theory testing (Shmueli, 2010). Moreover, the sole emphasis is

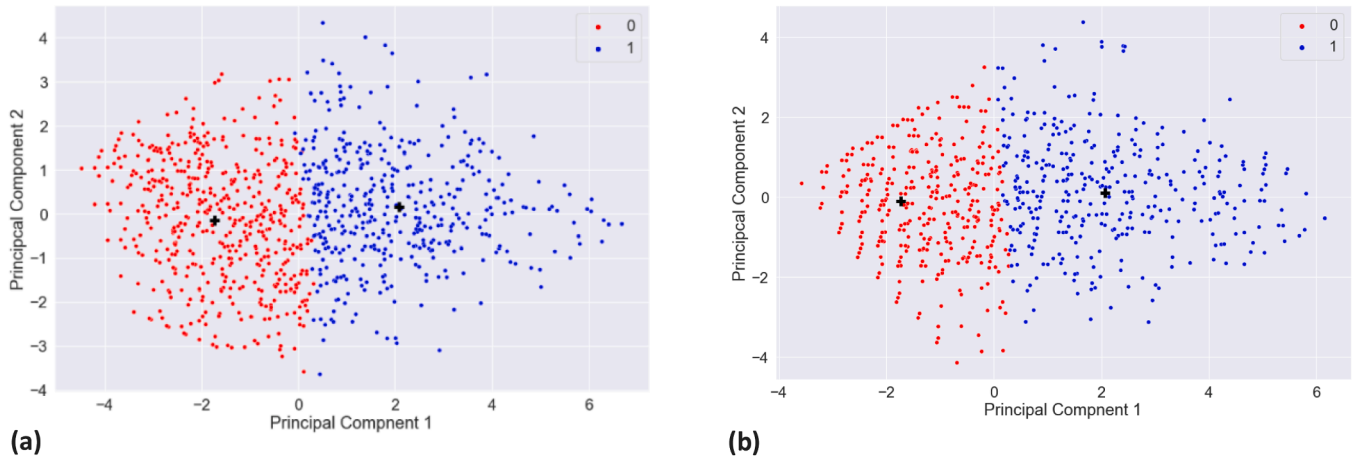

Fig. 12. Clusters for (a) T data set of smartphone security, (b) S data set of smartphone security.

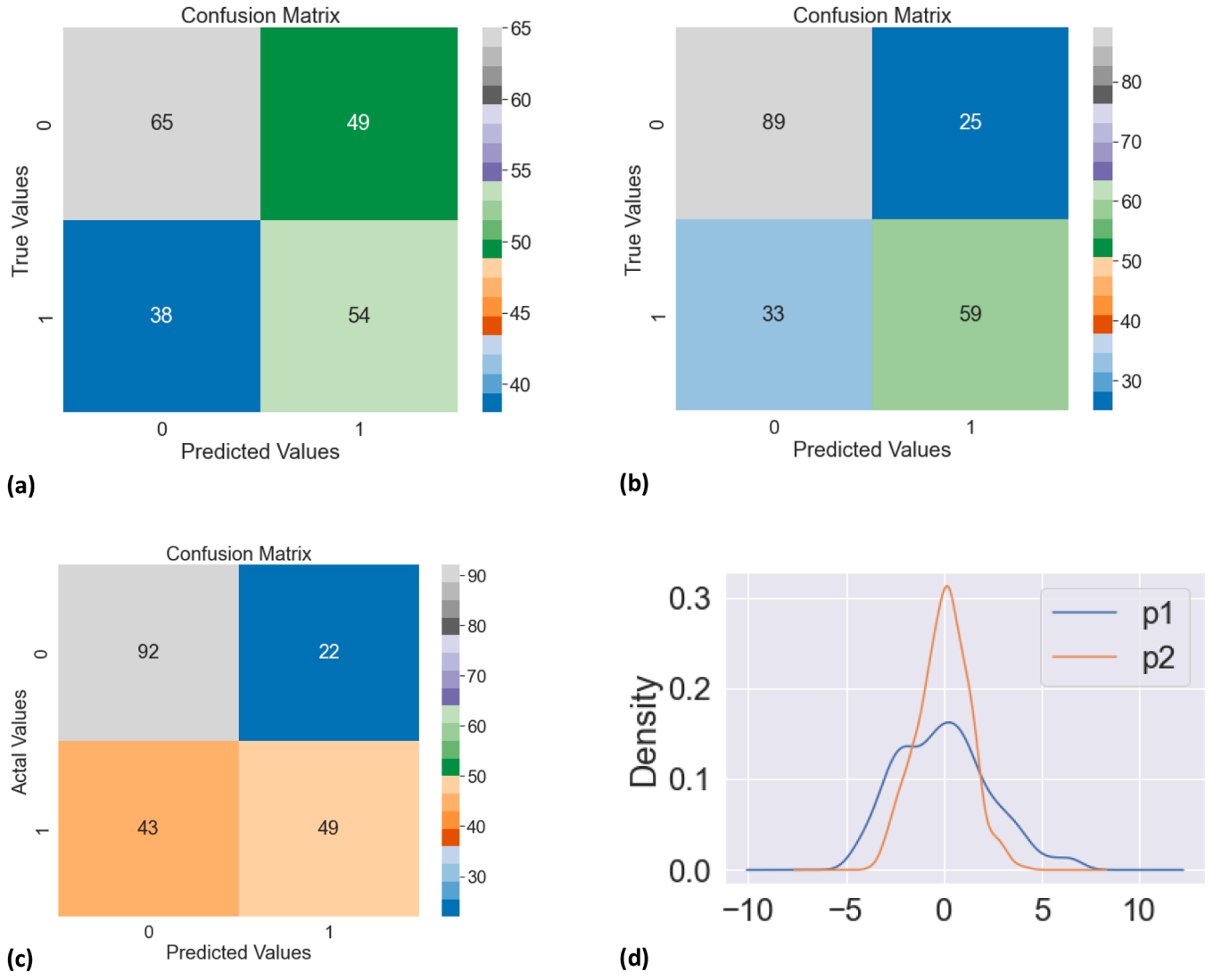

Fig. 13. Confusion matrices for technical dimension of smartphone security (a) DT, (b) KNN, (c) SVM; (d) PDG principal components of Technical dimension of smartphone security.

on accuracy and high values of this measure alone does not necessarily means theory being tested is a valid theory. That is to say the two approaches complement each other by offering insights that are present in one. The explanatory modeling allows to understand the relationship between variables while predictive modeling help in providing the predictive accuracy of the models. In this study, the employment of SEM

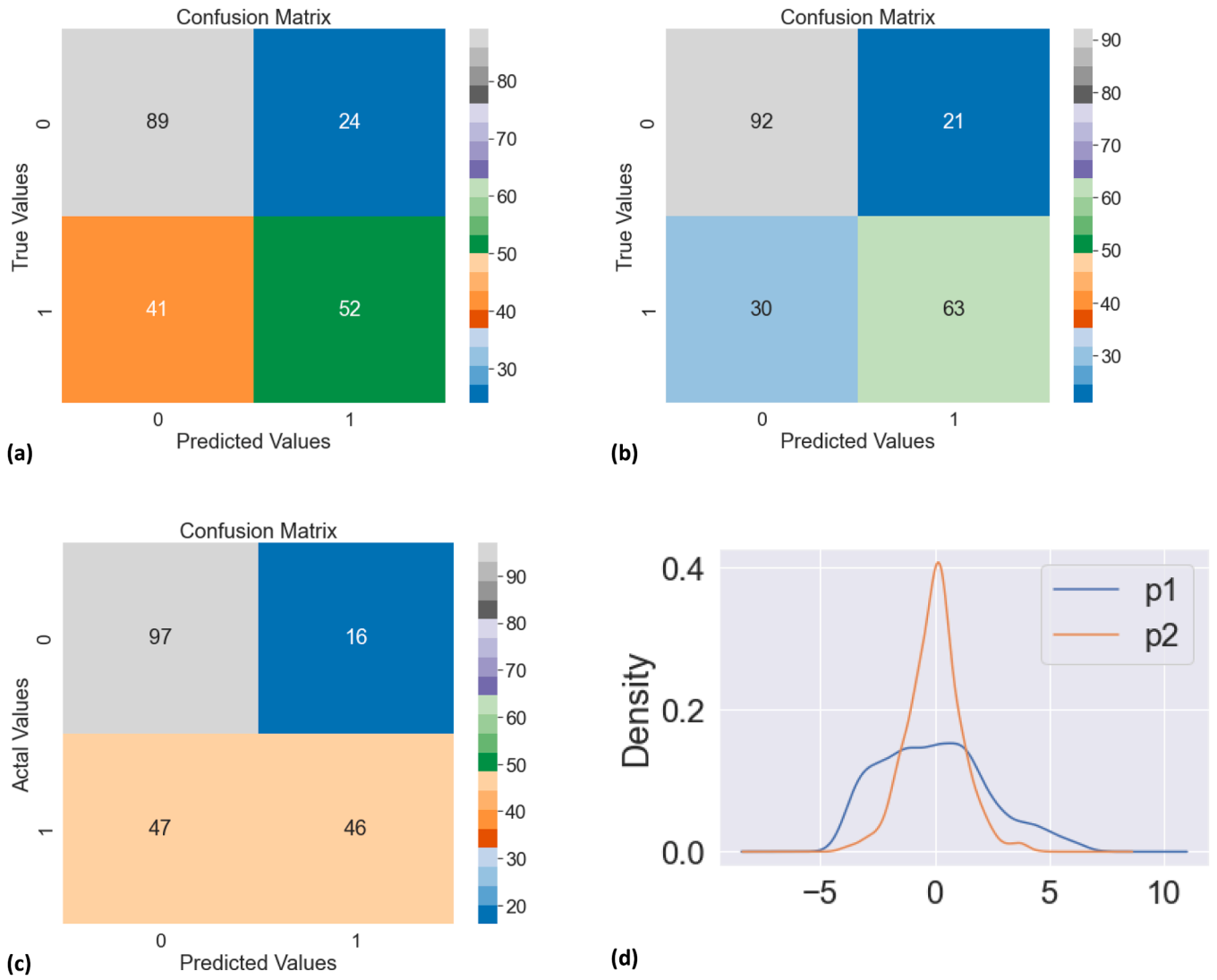

Fig. 14. Confusion matrices for S dimension of smartphone security (a) DT, (b) KNN, (c) SVM; (d) PDG principal components of Social dimension of smartphone security.

and ML algorithms have validated the PMT for Cybersecurity behaviors. While exclusively the explanatory modeling contributes towards the understanding of the cybersecurity behaviors, these claims have been evaluated in terms of predictive modeling. The complementarity of the two modeling approaches has given a more holistic understanding of the construct under study. Not only the data reflect the consistency with the theory as was gleaned from SEM, but the accuracy of around 63 % (a number not impressive in prediction) showed limitations in our understanding of cybersecurity PMT models. That is to say, forcing prediction of the explanatory model has revealed that it explains less than expected as reasoned in (Salganik et al., 2020; Ward et al., 2010) and opens avenues for findings and more complete explanations as argued in (Fudenberg et al., 2019).

The meta-level statistical analysis of cyber security behavior studies depicts that the *self-efficacy* is consistent in explanations of these models which was also found in this study's results. However, the predictive modeling revealed other PMT elements – *threat severity, response cost*, and *threat vulnerability* – to be important along with *self-efficacy* with average predictive accuracy above 60 % for both computer and smartphone security. As reasoned (Hofman et al., 2021), this predictive accuracy has set the baseline for cybersecurity studies employing PMT – i.e. any additional factors in the model is given importance not merely on its absolute performance but should be compared with the baseline accuracy of this study. That is to say, simply focusing on the increase in variance in performance with additional factors, the predictive accuracy benchmark of this study has to be taken into account.

6.3. *Implications for research*

The study has multiple key takeaways;

1. We have shown the complementarity of the explanatory and predictive modeling approaches to indicate similar findings in terms of PMTs antecedents. The distinctiveness of the predictive modeling has indicated predictive accuracy benchmarks for computer and smartphone security behaviors.
2. We have augmented SEM with predictive modeling to get the predictive power of PMT model in the answer to research call of (Alassaf and Alkhalifah, 2021).
3. We have made use of three ML algorithms for analysis thereby addressing the biases of cross-sectional research designs.
4. We have empirically tested the predictive power of PMT for two cyber security behaviors – computer security and smart phone security.

5. We have elaborated on micro and macro predictive modeling for security behaviors (smartphone and computer) to get a more nuanced insights.

A number of research implications can be garnered from this study. First of all, this is one of the first studies that present the explanatory and predictive power of PMT for cybersecurity behaviors and hence offers a predictive benchmark for further studies. The identification of non-significant antecedents (found in SEM analysis) as important features in predictive analysis allows for understanding cybersecurity behaviors without the assumptions of linearity relationship between data. Moreover, the identification of the most important antecedents in predicting cybersecurity behaviors has given ample evidence for improving cybersecurity training and education. The predictive modeling results imply that the predictive accuracy benchmark of PMT for cyber security behaviors is approximately 70 % which is a fair threshold. As has been argued (Shmueli, 2010) that higher accuracy benchmarks depict the sufficiency of the theoretical model, 70 % cyber security behavior accuracy shows PMT to be an appropriate theory in understanding smartphone security behavior of the university-going students. Nevertheless the predictive benchmark of 70 % calls for future studies by interested researchers to consider other factors such as fear appeal (Boss et al., 2015) and testing of different nomology of PMT models (Haag et al., 2021). The explanatory modeling complemented by predictive modeling has implications for PMT theory's validation in which the internal mechanism of interplay between the antecedent variables and the dependent variables has been explained and validation has been done via future outcome prediction.

6.4. Practical implications

From a practical point of view, the results of this study show the suitability of protection motivation theory in development of cybersecurity educational programs in higher education institutes. The dimension wise predictive analysis allowed for identification of critical drivers of overall security performance. Notably, proactive awareness emerged as the most influential factor in the all-inclusive clustering. This can be used in future focused analyses to measure the polarity or consistency of security behaviors, helping design more effective interventions. From PMT's perspective, the cybersecurity educational programs should emphasize on *self-efficacy, threat severity* and *response efficacy* related content of computer and smartphone security devices. Special care should be given to the participants who do not have IT-related background. Such students should be briefed on various security controls available on computer and smartphone devices and on the knowledge that how enabling security controls will build their smartphone *self-efficacy*. To cater for the evolving cyber threat landscape and the evolving nature of computer and security devices, the training content should be regularly updated and imparted to the target participants. Doing so will result in a cybersecurity eco system that will maintain safe and secure usage of the cyberspace for the tertiary institutes.

6.5. Limitations and future research directions

There are a number of limitations in this study. The use of self-report data may constitute social desirability biases. However, the provision of assurance of confidentiality and anonymity by the researchers may minimize situational and dispositional characteristic biases. The generalizability of the findings is another limitation which is limited to students' population. The students are the early adopter of the technology and are frequent users of smartphones and computers therefore, the results are generalizable to that population. Surveys based on student populations are from a narrow age range and may not generalize to other demographic groups, such as older adults and working professionals. Similarly, students are typically from same socioeconomic backgrounds and hence the perspectives from different socioeconomic classes is limited in this study. Yet another concern is the change of habits over time for students' population and the results of this study may not extend over time. Nevertheless, we plan to gather more data from faculty, staff and IT staff specifically digital natives from universities in the future. The socioeconomic background of the participants is also been reported to have influence on the cybersecurity behaviors (Mohammad et al., 2022). Therefore, we also plan to incorporate socioeconomic and digital divide variables in the future studies in the same vein as that of (Khan et al., 2023; 2022) which is key in understanding cybersecurity for non-WEIRD populations. Further extending the variables in the PMT models, the future studies may also include cross cultural populations to understand the explanatory and predictive power with respect to collectivism and power distance dimensions of Hofsted model (Hofstede, 2011). The elements of PMT that were considered in the models of this study are *threat appraisal* and *coping appraisal*. In the future studies, it would be important to incorporate fear appeals to test the PMT model for adaptive and maladaptive behaviors for computer security and smartphone security. From machine learning perspective, the future studies will incorporate the use of Pareto principle before clustering that will yield the most important factors thereby improving the clustering outcomes by making them distinct.

6.6. Conclusion

This study reported on explanatory modeling via SEM and predictive modeling via ML for protection motivation theoretical models of computer security and smartphone security. Three ML algorithms – DT, SVM and KNN were employed to ascertain the predictive accuracy of the two models. Furthermore, different types of computer security behaviors and smartphone security behaviors also underwent ML analysis to find the predictive power of PMT for these specific behaviors. Wrapper feature selection approach was also employed to find out the elements of PMT which were the most important predictors of cybersecurity behaviors. The SEM results revealed that perceived severity, *self-efficacy* and *response efficacy* influence computer and smartphone security behaviors which is consistent with the prior literature reviewed. The intention to secure devices is also positively associated with cyber security behaviors. The predictive accuracy for computer security PMT model is 73 % while that for smartphone security PMT model is 68 %. For specific computer security behaviors of device security, proactive awareness, password generation and updating are 78 %, 69 %, 75 % and 73 % respectively. The predictive accuracy for Technical smartphone security behavior was 71 % while that for social smartphone security behavior was 75 %. The highest accuracy was achieved by the KNN algorithm. The most important features for computer and smartphone security behavior as found by the wrapper feature selection method were *self-efficacy, response efficacy* and intention to secure device which were also consistent with the previous results found for explanatory modeling. The predictive modeling results validated the SEM results for PMT models of computer and smartphone security behaviors.

CRediT authorship contribution statement

Uzma Kiran: Visualization, Formal analysis, Data curation. **Naurin Farooq Khan:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Conceptualization. **Hajra Murtaza:** Writing – review & editing, Writing – original draft, Visualization, Investigation, Formal analysis. **Ali Farooq:** Writing – review & editing, Writing – original draft, Methodology, Conceptualization. **Henri Pirkkalainen:** Writing – review & editing, Methodology, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

The corresponding author has the conflict of interest with: 1) Steven Furnell, working on a same project 2) Nathan Clarke, by connection of Steven Furnell who are co-EICs of Computers & Information Security journal 2) Kim-Kwang Raymond Choo, other reasons If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix

Construct	Items
Perceived Severity (Thompson et al., 2017)	PS1. A security breach on my device would be a serious problem for me.
	PS2. A loss of information resulting from hacking would be a serious problem for me
	PS3. Having my confidential information on my device accessed by someone without my consent or knowledge would be a serious problem for me.
	PS4. Having someone successfully attack and damage my device would be very problematic for me.
	PS5. I view information security attacks on me as harmful.
	PS6. I believe that protecting the information on my device is important.
Perceived vulnerability (Thompson et al., 2017)	PV1. I could be subject to a serious information security threat.
	PV2. I am facing more and more information security threats.
	PV3. I feel that my device could be vulnerable to a security threat.
	PV4. It is likely that my device will be compromised in the future.
	PV5. My information and data is vulnerable to security breaches.
	PV6. I could fall victim to a malicious attack if I fail to follow good security practices.
Response cost (Thompson et al., 2017)	RC1. Taking security measures inconveniences me.
	RC2. There are too many overheads associated with taking security measures to protect my device.
	RC3. Taking security measures would require considerable investment of effort.
	RC4. Implementing security measures on my device would be time consuming.
	RC5. The cost of implementing recommended security measures exceeds the benefits.
	RC6. The impact of security measures on my productivity exceeds the benefits.
Response efficacy (Thompson et al., 2017)	RE1. Enabling security measures on my device will prevent security breaches.
	RE2. Implementing security measures on my device is an effective way to prevent hackers.
	RE3. Enabling security measures on my device will prevent hackers from stealing my identity.
	RE4. The preventative measures available to stop people from getting confidential personal or financial information on my device are effective.
Self-efficacy (Thompson et al., 2017)	SE1. I feel comfortable taking measures to secure my device.
	SE2. Taking the necessary security measures is directly under my control.
	SE3. I have the resources and the knowledge to take the necessary security measures.
	SE4. Taking the necessary security measures in easy.
	SE5. I can protect my device by myself, I can enable security measures on my device.
Security Intentions (Thompson et al., 2017)	SI1. I am likely to take security measures on my devices (smartphone, computer/laptop).
	SI2. I will take security measures to protect my devices (smartphone, computer/laptop).
	SI3. It is my intention to take measures to protect my devices (smartphone, computer/laptop).
Smartphone Security Behavior (Huang et al., 2020)	SS1. I reset my Advertising ID on my smartphone.
	SS2. I hide device in my smartphone's Bluetooth settings.
	SS3. I change my passcode/PIN for my smartphone's screen lock at a regular basis.
	SS4. I manually cover my smartphone's screen when using it in the public area (e.g., bus or subway).
	SS5. I use an adblocker on my smartphone.
	SS6. I use an anti-virus app.
	SS7. I use a Virtual Private Network (VPN) app while connected to a public network.
	SS8. I turn off WiFi on my smartphone when not actively using it.
	SS9. I care about the source of the app when performing financial and/or shopping tasks on that app.
	SS10. I take care of the source of the app when performing financial and / or purchasing work on this app.
	SS11. When downloading an app, I check that the app is from the official/expected source.
	SS12. I verify the recipient/sender before sharing text messages or other information using smartphone apps.
	SS13. I delete any online communications (i.e. texts, emails, social media posts) that look suspicious.
	SS14. I pay attention to the pop-ups on my smartphone when connecting it to another device (e.g. laptop, desktop).
Computer Security Behavior (Egelman and Peer, 2015)	CS1. I set my computer screen to automatically lock if I don't use it for a prolonged period of time.
	CS2. I use a password/passcode to unlock my laptop or tablet.
	CS3. I manually lock my computer screen when I step away from it.
	CS4. I use a PIN or passcode to unlock my mobile phone.
	CS5. I change my passwords even if it is not needed.
	CS6. I use different passwords for different accounts that I have.
	CS7. When I create a new online account, I try to use a password that goes beyond the site's minimum requirements.
	CS8. I include special characters in my password even if it's not required.
	CS9. When someone sends me a link, I open it only after verifying where it goes.
	CS10. I know what website I'm visiting by looking at the URL bar, rather than by the website's look and feel.
	CS11. I verify that information will be sent securely (e.g. SSL, https://, a lock icon) before I submit it to websites.
	CS12. When browsing websites, I mouseover links to see where they go, before clicking them.
	CS13. If I discover a security problem, I fix or report it rather than assuming somebody else will.
	CS14. When I'm prompted about a software update, I install it right away.
	CS15. I try to make sure that the programs I use are up-to-date.
	CS16. I verify that my anti-virus software has been regularly updating itself.

Acronym List:

AI - Artificial intelligence
AFM - Absolute fit measure
AVE - Average variance extracted
BYOD - Bring your own device
CB-SEM - Covariance based structure equation modeling
CFI - Comparative fit index
CR - Composite reliability
CSB - Computer security behavior
DS - Device security
DT - Decision Tree
fsQCA - Fuzzy-set qualitative comparative analysis
GDT - general deterrence theory
GFI - Goodness of fit index
IFM - Incremental fit measure
KNN - K nearest neighbor
ML - Machine learning
NN - Neural network
PG - Password generation
PA - Proactive awareness
PFM - Parsimonious fit measure
PMT - Protection motivation theory
PNFI - Parsimony normed fit index
PCA – Principal component analysis
PCFI - Parsimonious comparative fit index
PDG - Probability density graph
RC - Perceived response cost
RE - Perceived response efficacy
RMSEA - Root mean square error
SE - Perceived self-efficacy
SeBIS - Security behavior intention scale
SI - Intention to secure computer device
SSB - Smartphone security behavior
SVM - Support vector machine
TLI - Tucker-lewis index
TPB - Theory of planned behavior
TS - Perceived threat severity
TV - Perceived threat vulnerability
UP - Updating
WEIRD - Western, educated, industrialized, rich and democratic

Data availability

Data will be made available on request.

References

Alassaf, M., Alkhalifah, A., 2021. Exploring the Influence of Direct and Indirect Factors on Information Security Policy Compliance: A Systematic Literature Review. IEEE Access.

Almazroi, A.A., Mohamed, O.A., Shamim, A., Ahsan, M., 2020. Evaluation of state-of-the-art classifiers: A comparative study. Res.p. J. Comput. 1 (1), 22–29.

Almheiri, E., Al-Emran, M., Al-Sharafi, M.A., Arpaci, I., 2024. Drivers of smartwatch use and its effect on environmental sustainability: Evidence from SEM-ANN approach. Asia-Pacific J. Bus. Adm.

Alshurideh, M., Al Kurdi, B., Salloum, S.A., Arpaci, I., Al-Emran, M., 2023. Predicting the actual use of m-learning systems: A comparative approach using PLS-SEM and machine learning algorithms. Interact. Learn. Environ. 31 (3), 1214–1228.

Alwabel, A.S.A., Zeng, X.-J., 2021. Data-driven modeling of technology acceptance: A machine learning perspective. Expert. Syst. Appl. 185, 115584.

Ameen, N., Tarhini, A., Shah, M.H., Madichie, N., Paul, J., Choudrie, J., 2021. Keeping customers' data secure: A cross-cultural study of cybersecurity compliance among the Gen-Mobile workforce. Comput. Human. Behav. 114, 106531.

Anderson, J.C., Gerbing, D.W., 1988. Structural equation modeling in practice: A review and recommended two-step approach. Psychol. Bull. 103 (3), 411.

Arnett, J.J., 1996. Sensation seeking, aggressiveness, and adolescent reckless behavior. Pers. Individ. Dif. 20 (6), 693–702.

Arpaci, I., 2023. Predictors of financial sustainability for cryptocurrencies: An empirical study using a hybrid SEM-ANN approach. Technol. Forecast. Soc. Change 196, 122858.

Bahari, M., Arpaci, I., Azmi, N.F.M., Shuib, L., 2023. Predicting the intention to use learning analytics for academic advising in higher education. Sustainability 15 (21), 15190.

Boss, S.R., Galletta, D.F., Lowry, P.B., Moody, G.D., Polak, P., 2015. What do systems users have to fear? Using fear appeals to engender threats and fear that motivate protective security behaviors. MIS Q. 39 (4), 837–864.

Breitinger, F., Tully-Doyle, R., Hassenfeldt, C., 2020. A survey on smartphone user's security choices, awareness and education. Comput. Secur. 88, 101647.

Brodin, M., Rose, J., 2020. Mobile information security management for small organisation technology upgrades: The policy-driven approach and the evolving implementation approach. Int. J. Mobile Commun. 18 (5), 598–618.

Butler, R., 2020. A systematic literature review of the factors affecting smartphone user threat avoidance behaviour. Inf. Comput. Secur. 28 (4), 555–574.

Byrne, B.M., 2013. Structural Equation Modeling with EQS: Basic Concepts, Applications, and Programming. Routledge.

Cheng, S.-I., 2011. Comparisons of competing models between attitudinal loyalty and behavioral loyalty. Int. J. Bus. Soc. Sci. 2 (10), 149–166.

Chin, A.G., Little, P., Jones, B.H., 2020. An analysis of smartphone security practices among undergraduate business students at a regional public university. Int. J. Educ. Dev. Using. Inf. Commun. Technol. 16 (1), 44–61.

Crossler, R., Bélanger, F., 2014. An extended perspective on individual security behaviors: Protection motivation theory and a unified security practices (USP) instrument. ACM SIGMIS Database: DATABASE Adv. Inf. Syst. 45 (4), 4. Article.

Crossler, R.E., Johnston, A.C., Lowry, P.B., Hu, Q., Warkentin, M., Baskerville, R., 2013. Future directions for behavioral information security research. Comput. Secur. 32, 90–101. https://doi.org/10.1016/j.cose.2012.09.010.

Crossler, R.E., Long, J.H., Loraas, T.M., Trinkle, B.S., 2014. Understanding compliance with bring your own device policies utilizing protection motivation theory: Bridging the intention-behavior gap. J. Inf. Systems 28 (1), 209–226.

Dahabiyeh, L., Farooq, A., Ahmad, F., & Javed, Y. (2023). Explaining technology migration against the change in terms of use: An fsQCA approach. *Information Technology & People, ahead-of-print*(ahead-of-print). https://doi.org/10.1108/ITP-07-2022-0498.

Dang-Pham, D., Pittayachawan, S., 2015. Comparing intention to avoid malware across contexts in a BYOD-enabled Australian university: a protection motivation theory approach. Comput. Secur. 48, 281–297.

Das, A., Khan, H.U., 2016. Security behaviors of smartphone users. Inf. Comput. Secur.

Dawie, F.J., Masrek, M.N., Rahman, S.A., 2022. Systematic Literature Review: Information security behaviour on smartphone users. Environ.-Behav. Proc. J. 7, 275–281. SI10.

Doane, A.N., Boothe, L.G., Pearson, M.R., Kelley, M.L., 2016. Risky electronic communication behaviors and cyberbullying victimization: an application of protection motivation theory. Comput. Human. Behav. 60, 508–513.

Dubin, R., 1969. Theory Building. The Free Press.

Egelman, S., Harbach, M., Peer, E., 2016. Behavior ever follows intention. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: CHI'16, pp. 1–5.

Egelman, S., Peer, E., 2015. Scaling the security wall: Developing a security behavior intentions scale (sebis). In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, pp. 2873–2882.

Fan, Y., Chen, J., Shirkey, G., John, R., Wu, S.R., Park, H., Shao, C., 2016. Applications of structural equation modeling (SEM) in ecological studies: An updated review. Ecol. Process. 5, 1–12.

Farooq, A., Kakakhel, S., Virtanen, S., Isoaho, J., 2015. A taxonomy of perceived information security and privacy threats among IT security students. In: 2015 10th International *Conference for Internet Technology and Secured Transactions (ICITST)*, pp. 280–286. https://doi.org/10.1109/ICITST.2015.7412106.

Farooq, A., Laato, S., Islam, A.N., 2020. Impact of online information on self-isolation intention during the COVID-19 pandemic: cross-sectional study. J. Med. Internet. Res. 22 (5), e19128.

Farooq, A., Ndiege, J.R.A., Isoaho, J., 2019. Factors affecting security behavior of Kenyan students: An integration of protection motivation theory and theory of planned behavior. In: 2019 IEEE AFRICON, pp. 1–8.

Floyd, D.L., Prentice-Dunn, S., Rogers, R.W., 2000. A meta-analysis of research on protection motivation theory. J. Appl. Soc. Psychol. 30 (2), 407–429.

Fornell, C., Larcker, D.F., 1981. Evaluating structural equation models with unobservable variables and measurement error. J. Mark. Res. 18 (1), 39–50. https://doi.org/10.1177/002224378101800104.

Forster, M., Sober, E., 1994. How to tell when simpler, more unified, or less ad hoc theories will provide more accurate predictions. Br. J. Philos. Sci. 45 (1), 1–35.

Fudenberg, D., Kleinberg, J., Liang, A., & Mullainathan, S. (2019). Measuring the completeness of theories. *arXiv Preprint arXiv:1910.07022*.

Geisser, S., 1993. An introduction to predictive inference. Chapman and Hall, New York.

Giwah, A.D., Wang, L., Levy, Y., Hur, I., 2019. Empirical assessment of mobile device users' information security behavior towards data breach: Leveraging protection motivation theory. J. Intell. Cap.

Haag, S., Siponen, M., Liu, F., 2021. Protection Motivation Theory in Information Systems Security Research: A Review of the Past and a Road Map for the Future. ACM SIGMIS Database: DATABASE for Adv. Inf. Syst. 52 (2), 25–67.

Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E., Tatham, R., 2010. Multivariate Data Analysis: Pearson Education. Pretince Hall, New Jersey.

Hamka, M., Ramdhoni, N., 2022. K-means cluster optimization for potentiality student grouping using elbow method. In: AIP Conference Proceedings, p. 2578. https://pubs.aip.org/aip/acp/article-abstract/2578/1/060011/2830069.

Harris, M.A., Furnell, S., Patten, K., 2014. Comparing the mobile device security behavior of college students and information technology professionals. J. Inf. Privacy Secur. 10 (4), 186–202.
Hina, S., Selvam, D.D.D.P., Lowry, P.B., 2019. Institutional governance and protection motivation: Theoretical insights into shaping employees' security compliance behavior in higher education institutions in the developing world. Comput. Secur. 87, 101594.
Hofman, J.M., Watts, D.J., Athey, S., Garip, F., Griffiths, T.L., Kleinberg, J., Margetts, H., Mullainathan, S., Salganik, M.J., Vazire, S., 2021. Integrating explanation and prediction in computational social science. Nature 595 (7866), 181–188.
Hofstede, G., 2011. Dimensionalizing cultures: The Hofstede model in context. Online Read. Psychol. Cult. 2 (1), 8.
Hovav, A., Putri, F.F., 2016. This is my device! Why should I follow your rules? Employees' compliance with BYOD security policy. Pervasive Mob. Comput. 32, 35–49.
Huang, H.-Y., Demetriou, S., Banerjee, R., Tuncay, G. S., Gunter, C. A., & Bashir, M. (2020). Smartphone Security Behavioral Scale: A New Psychometric Measurement for Smartphone Security. *arXiv Preprint arXiv:2007.01721*.
Ifinedo, P., 2012. Understanding information systems security policy compliance: An integration of the theory of planned behavior and the protection motivation theory. Comput. Secur. 31 (1), 83–95.
Johnston, A.C., Warkentin, M., 2010. Fear appeals and information security behaviors: an empirical study. MIS Q. 549–566.
Jolliffe, I., 2005. Principal Component Analysis. In: Everitt, B.S., Howell, D.C. (Eds.), Encyclopedia of Statistics in Behavioral Science, 1st ed. Wiley. https://doi.org/10.1002/0470013192.bsa501.
Jones, B.H., Chin, A.G., 2015. On the efficacy of smartphone security: A critical analysis of modifications in business students' practices over time. Int. J. Inf. Manage 35 (5), 561–571.
Khan, N.F., Ikram, N., Murtaza, H., Asadi, M.A., 2021. Social media users and cybersecurity awareness: Predicting self-disclosure using a hybrid artificial intelligence approach. Kybernetes 52 (1), 401–421. https://doi.org/10.1108/K-05-2021-0377.
Khan, N.F., Ikram, N., Saleem, S., 2023. Digital divide and socio-economic differences in smartphone information security behaviour among university students: Empirical evidence from Pakistan. Int. J. Mobile Commun. 1. https://doi.org/10.1504/ijmc.2023.10042359.
Khan, N.F., Ikram, N., Saleem, S., Zafar, S., 2022. Cyber-security and risky behaviors in a developing country context: A Pakistani perspective. Secur. J. https://doi.org/10.1057/s41284-022-00343-4.
Khan, N.F., Yaqoob, A., Khan, M.S., Ikram, N., 2022. The cybersecurity behavioral research: a tertiary study. Comput. Secur. 120, 102826.
Kline, R.B., 2015. Principles and Practice of Structural Equation Modeling. Guilford Publications.
Knapova, L., Kruzikova, A., Dedkova, L., Smahel, D., 2021. Who Is smart with their smartphones? Determinants of smartphone security behavior. Cyberpsychol. Behav. Soc. Network. 24 (9), 584–592.
Kohavi, R., John, G.H., 1998. The wrapper approach. In: Liu, H., Motoda, H. (Eds.), Feature Extraction, Construction and Selection. Springer US, pp. 33–50. https://doi.org/10.1007/978-1-4615-5725-8_3.
Lebek, B., Uffen, J., Neumann, M., Hohler, B., Breitner, M.H., 2014. Information Security Awareness and Behavior: A Theory-Based Literature Review. Management Research Review.
Li, L., He, W., Xu, L., Ash, I., Anwar, M., Yuan, X., 2019. Investigating the impact of cybersecurity policy awareness on employees' cybersecurity behavior. Int. J. Inf. Manage 45, 13–24.
Li, L., Xu, L., He, W., 2022. The effects of antecedents and mediating factors on cybersecurity protection behavior. Comput. Hum. Behav. Rep. 5, 100165.
Li, Y., Xin, T., Siponen, M., 2021. Citizens' Cybersecurity Behavior: Some Major Challenges. IEEE Security & Privacy.
Liébana-Cabanillas, F., Lara-Rubio, J., 2017. Predictive and explanatory modeling regarding adoption of mobile payment systems. Technol. Forecast. Soc. Change 120, 32–40. https://doi.org/10.1016/j.techfore.2017.04.002.
Ling, C.X., Huang, J., Zhang, H., 2003. AUC: A better measure than accuracy in comparing learning algorithms. In: Conference of the Canadian Society for Computational Studies of Intelligence, pp. 329–341.
Luuk, B., Remco, S., Rutger, L.E., 2023. Protecting your business against ransomware attacks? Explaining the motivations of entrepreneurs to take future protective measures against cybercrimes using an extended protection motivation theory model. Comput. Secur., 103099
Mai, P.T., Tick, A., 2021. Cyber security awareness and behavior of youth in smartphone usage: A comparative study between university students in Hungary and Vietnam. Acta Polytech. Hung. 18, 67–89.
Maier, C., Thatcher, J.B., Grover, V., Dwivedi, Y.K., 2023. Cross-sectional research: A critical perspective, use cases, and recommendations for IS research. Int. J. Inf. Manage 70, 102625. https://doi.org/10.1016/j.ijinfomgt.2023.102625.
Mills, A., & Sahi, N. (2019). *An empirical study of home user intentions towards computer security*.
Mithas, S., Xue, L., Huang, N., Burton-Jones, A., 2022. Editor's Comments: Causality Meets Diversity in Information Systems Research. Manage. Inf. Syst. Q. 46 (3) iii–xviii.
Mohammad, T., Hussin, N.A.M., Husin, M.H., 2022. Online safety awareness and human factors: An application of the theory of human ecology. Technol. Soc. 68, 101823.
Moody, G.D., Siponen, M., Pahnila, S., 2018. Toward a unified model of information security policy compliance. MIS Q. 42 (1).
Mou, J., Cohen, J.F., Bhattacherjee, A., Kim, J., 2022. A test of protection motivation theory in the information security literature: a meta-analytic structural equation modeling approach. J. Assoc. Inf. Syst. 23 (1), 196–236.
Mou, J., Cohen, J., Kim, J., 2017. A meta-analytic structural equation modeling test of protection motivation theory in information security literature. In: International Conference on Information Systems.
Murphy, K.P., 2012. Machine Learning: A Probabilistic Perspective. MIT press.
Ng, B.-Y., Kankanhalli, A., Xu, Y.C., 2009. Studying users' computer security behavior: A health belief perspective. Decis. Support. Syst. 46 (4), 815–825.
Nowrin, S., Bawden, D., 2018. Information security behaviour of smartphone users. Inf. Learn. Sci.
Ogbanufe, O., Crossler, R.E., Biros, D., 2023. The valued coexistence of protection motivation and stewardship in information security behaviors. Comput. Secur. 124, 102960.
Osisanwo, F.Y., Akinsola, J.E.T., Awodele, O., Hinmikaiye, J.O., Olakanmi, O., Akinjobi, J., 2017. Supervised machine learning algorithms: Classification and comparison. Int. J. Comput. Trends Technol. (IJCTT) 48 (3), 128–138.
Palanisamy, R., Norman, A.A., Kiah, M.L.M., 2020. Compliance with bring your own device security policies in organizations: a systematic literature review. Comput. Secur., 101998
Panovska-Griffiths, J., Kerr, C.C., Waites, W., Stuart, R.M., 2021. Mathematical modeling as a tool for policy decision making: Applications to the COVID-19 pandemic. In: Handbook of Statistics, 44. Elsevier, pp. 291–326.
Posey, C., Roberts, T.L., Lowry, P.B., 2015. The impact of organizational commitment on insiders' motivation to protect organizational information assets. J. Manage. Inf. Syst. 32 (4), 179–214.
Posey, C., Roberts, T., Lowry, P.B., Courtney, J., Bennett, B., 2011. Motivating the insider to protect organizational information assets: Evidence from protection motivation theory and rival explanations. In: The Dewald Roode Workshop in Information Systems Security, pp. 22–23.
Rajab, M., Eydgahi, A., 2019. Evaluating the explanatory power of theoretical frameworks on intention to comply with information security policies in higher education. Comput. Secur. 80, 211–223.
Rogers, R.W., 1975. A protection motivation theory of fear appeals and attitude change1. J. Psychol. 91 (1), 93–114.
Rogers, R.W., 1983. Cognitive and psychological processes in fear appeals and attitude change: A revised theory of protection motivation. In: Society Psychophysiology: A Sourcebook, pp. 153–176.
Rogers, R.W., Prentice-Dunn, S., 1997. *Protection motivation theory* (Handbook of Health Behavior Research 1: Personal and social determinants). Plenum press.
Salganik, M.J., Lundberg, I., Kindel, A.T., Ahearn, C.E., Al-Ghoneim, K., Almaatouq, A., Altschul, D.M., Brand, J.E., Carnegie, N.B., Compton, R.J., 2020. Measuring the predictability of life outcomes with a scientific mass collaboration. Proc. Natl. Acad. Sci. 117 (15), 8398–8403.
Sarstedt, M., Hair, J.F., Ringle, C.M., Thiele, K.O., Gudergan, S.P., 2016. Estimation issues with PLS and CBSEM: Where the bias lies! J. Bus. Res. 69 (10), 3998–4010.
Schneier, B., 2015. Secrets and Lies: Digital Security in a Networked World. John Wiley & Sons.
Shah, P., Agarwal, A., 2020. Cybersecurity behaviour of smartphone users in India: An empirical analysis. Inf. Comput. Secur.
Sharma, S., Aparicio, E., 2022. Organizational and team culture as antecedents of protection motivation among IT employees. Comput. Secur. 120, 102774.
Shmueli, G., 2010. To explain or to predict? Statist. Sci. 25 (3), 289–310.
Shmueli, G., Koppius, O.R., 2011. Predictive analytics in information systems research. MIS Q. 553–572.
Siponen, M., Rönkkö, M., Fufan, L., Haag, S., Laatikainen, G., 2024. Protection motivation theory in information security behavior research: reconsidering the fundamentals. Commun. Assoc. Inf. Syst. 53 (1), 1136–1165. https://doi.org/10.17705/1CAIS.05348.
Sommestad, T., Karlzén, H., Hallberg, J., 2015a. A meta-analysis of studies on protection motivation theory and information security behaviour. Int. J. Inf. Secur. Privacy (IJISP) 9 (1), 26–46.
Sommestad, T., Karlzén, H., Hallberg, J., 2015b. A meta-analysis of studies on protection motivation theory and information security behaviour. Int. J. Inf. Secur. Privacy (IJISP) 9 (1), 26–46.
Stylios, I., Kokolakis, S., Thanou, O., Chatzis, S., 2016. Users' attitudes on mobile devices: can users' practices protect their sensitive data?. In: MCIS 2016 Proceedings.
Sun, Y., Wang, N., Shen, X.-L., 2020. Toward a Configurational Protection Motivation Theory. In: *53rd Hawaii International Conference on System Sciences*. HICSS. https://scholarspace.manoa.hawaii.edu/items/a618cbee-c90c-4d72-91e7-ad1aaa0f8e6e.
Taber, K.S., 2018. The use of Cronbach's alpha when developing and reporting research instruments in science education. Res. Sci. Educ. 48 (6), 1273–1296.
Tarhini, A., AlHinai, M., Al-Busaidi, A.S., Govindaluri, S.M., Shaqsi, J.A., 2024. What drives the adoption of mobile learning services among college students: An application of SEM-neural network modeling. Int. J. Inf. Manag. Data Insights 4 (1), 100235. https://doi.org/10.1016/j.jjimei.2024.100235.
Thompson, N., McGill, T.J., Wang, X., 2017. Security begins at home": Determinants of home computer and mobile device security behavior. Comput. Secur. 70, 376–391.
Tsai, H.S., Jiang, M., Alhabash, S., LaRose, R., Rifon, N.J., Cotten, S.R., 2016. Understanding online safety behaviors: A protection motivation theory perspective. Comput. Secur. 59, 138–150.
Tu, C.Z., Adkins, J., Zhao, G.Y., 2019. Complying with BYOD security policies: A moderation model based on protection motivation theory. J. Midwest Assoc. Inf. Syst. (JMWAIS) 1, 11–28.

Vance, A., Siponen, M., Pahnila, S., 2012. Motivating IS security compliance: Insights from habit and protection motivation theory. Inf. Manage. 49 (3–4), 190–198.
Verkijika, S.F., 2018. Understanding smartphone security behaviors: An extension of the protection motivation theory with anticipated regret. Comput. Secur. 77, 860–870.
Vrhovec, S., Mihelič, A., 2021. Redefining threat appraisals of organizational insiders and exploring the moderating role of fear in cyberattack protection motivation. Comput. Secur. 106, 102309.
Ward, M.D., Greenhill, B.D., Bakke, K.M., 2010. The perils of policy by p-value: Predicting civil conflicts. J. Peace Res. 47 (4), 363–375.
Zhang, X.J., Li, Z., Deng, H., 2017. Information security behaviors of smartphone users in China: An empirical analysis. The Electronic Library.
Zhou, G., Gou, M., Gan, Y., Schwarzer, R., 2020. Risk awareness, self-efficacy, and social support predict secure smartphone usage. Front. Psychol. 11, 1066.