Journal of Informatics Education and Research
ISSN: 1526-4726
Vol 6 Issue 1 (2026)

Integrating Behavioural Theory and Machine Learning to Predict Life Insurance Purchase Intention in India

Dr. K. Sanjay,

Associate Professor, Dayananda Sagar University, Bangalore.

Abstract

Life insurance penetration In India, remains significantly below global averages. This despite a young population and rising income levels. Traditional theories, with its emphasis of life insurance being a rational product, have not been able to explain this paradox. It is important, therefore, to look beyond rationality and look at psychological and financial factors to explain this phenomenon. This study extended Theory of planned behaviour (TPB). This research integrated the psychological determinants of TPB i.e. attitude, subjective norms and PBC with financial drivers i.e. financial literacy, saving motive and risk aversion. Primary data was used in the study. The SEM model found that all the TPB constructs significantly predict intention with attitude being the strongest predictor. Financial literacy and Saving motive also had a significant and positive effect on intention. Additionally, the findings indicate that saving motivation serves as a mediating mechanism linking financial literacy to behavioural intention. Risk aversion showed a weaker but significant effect without moderation effect. To validate behavioural findings two ML models, using Random Forest and XGBoost were used. Both the ML models demonstrated superior predictive accuracy confirming nonlinear behavioural dynamics. The findings extend existing theoretical frameworks and offer meaningful implications for practitioners and policymakers for insurance policy design, marketing communication, and consumer inclusion strategies.

Keywords: Life insurance intention; Theory of Planned Behaviour; Financial literacy; Saving motive; Risk aversion; Machine learning

Introduction: India has a paucity of state-sponsored social security schemes. Therefore, life insurance has a critical role to play here. It provides financial succour to the grieving family who has lost their primary bread winner. In addition to this life insurance encourages long term savings. (Beck & Webb 2003) Economists have emphasised this dual role of this personal risk-protection financial instrument which is mitigating financial risk and encouraging long term savings. (Beck and Webb 2003). Despite the above logical reasons, not many people buy life insurance in India.

One of metrics that economists use to measure the depth of the life insurance market is penetration. This is the ratio of total life insurance premiums collected during the year to the GDP. Life insurance penetration in India is just 2.8% in FY2024 compared to the global average of 7%. (IRDA report 2024). This shows that purchase of life Insurance in India is much lower than their global peers. A study by Kakar & Shukla shows that only a quarter of Indian household own life insurance. The study also shows is that many of these same households, especially well-educated professionals, fully understand what a policy offers and its benefits. (Kakar & Sharma 2010). Despite these reasons these household are reluctant to purchase life insurance policies. Other studies too have reached the same conclusion. A recent study by Giri & Chatterjee 2021 suggest that many households never buy or don't continue their life insurance cover. A study by life insurance company SBI life found that only 6% of Indian consumers are adequately covered. (SBI Life Insurance & Deloitte, 2023)

The consequences of low insurance penetration are dire. Unexpected death of the primary breadwinner could leave the families saddled with unsustainable debt or they could in some cases slip into poverty. (Mahdzan & Victorian, 2013; Pauly et al., 2003). In addition to the financial risk, households also miss out on an opportunity for long term capital. Long term savings are essential to meet milestone expenses like meeting the cost of a child's higher education, retirement etc. Also, the economy of the country loses a source of long-term capital. (Beck & Webb, 2003; Liedtke, 2007).

The continued suboptimal uptake indicates that traditional economic theories, which assume rational purchase decisions, may be insufficient to explain the lack of adequate LI cover. There could be psychological factors which influence purchase of LI. Factors like risk perception, attitude, lack of trust, or limited product comprehension could have a meaningful role in shaping consumer decisions (Kahneman, 2011; Thaler & Sunstein, 2008, Ajzen 1991).

Life insurance studies in India have mostly focussed on macroeconomic parameters or socio-demographic factors like age, gender, income etc. These studies don't adequately explain why life insurance purchase is low in India. For e.g. two households, with similar age and education, could exhibit two very diverse behavioural pattern with respect to insurance. (Bhat & Jain, 2013; Dash & Mahapatra, 2014).

Recent studies have used psychological models like the Ajzen's behavioural decision-making framework more commonly referred to as Theory of Planned Behaviour (Ajzen, 1991). This research has extended the theoretical framework by incorporating financial motivations such as saving motives and risk aversion. Research findings have shown that attitudes, social influence, perceived control, saving motives, and risk aversion all have demonstrated a statistically meaningful positive influence on inclination to buy life insurance**.** Research shows that financial literacy has only a mediating effect. (Sanjay & Tewari, 2024; Sanjay et al., 2025). These findings suggest that the decision to buy life insurance is not purely rational. Psychological factors and financial motivators also strong effect.

Still, most of this research examines factors one at a time and depends heavily on self-reported survey data, which has its limits.

Reasons for sub-optimal purchase of life insurance have never been adequately explained. There is a clear and unexplained gap between predictions of rational economic theories and the life insurance purchased. (Browne & Kim, 1993). There is no comprehensive model explaining the interplay between psychological and financial motivators and its impact on an individual's inclination to purchase life insurance. Also, there is no study which has checked whether the model predicts behaviour. There has been no Indian study which combines behavioural analysis and the predictive power of machine-learning algorithms.

Most research studies in life insurance use traditional statistical tools like logistic regression or SEM. However, a basic requirement for these techniques is that the data is linear and normal. This limits their ability to predict complex decisions like purchase of life insurance. (Ajzen, 1991; Bhatia et al., 2021). So, the accuracy of their predictions is rather low when applied to complex buying decisions like purchase of life insurance.

ML models like Random Forest and XGBoost offer a lot of advantages for researchers. It can identify relative importance of factors in prediction. Also, it can identify nonlinear interactions which traditional regression sometimes miss. Recent studies show that ML models provide superior predictive performance when used in life insurance studies, with some studies showing 92% accuracy. (Ghafar et.al. 2025). Azzone et al. (2022) showed that Random Forest models significantly outperformed traditional logistic regression in predicting life-insurance lapses, highlighting the advantage of ensemble methods in capturing complex behavioural patterns. Similar findings are reported in studies that apply gradient-boosting techniques and decision trees to insurance datasets. ML models have consistently yielded higher accuracy than econometric approaches (Shamsuddin et al., 2023). While ML based tools have shown their utility, very few research studies have used these tools. This shows that there is immense potential to use these tools in research.

Another important reason to use ML is that it has demonstrated its ability to also consider variables which are difficult to model using traditional statistical tools. (Gramegna and Giudici 2020). Hence, explainability allows researchers to interpret complex ML models and derive importance into meaningful insights. This enables alignment with behavioural theories such as TPB (Gramegna & Giudici, 2020). Consequently, ML complements rather than replace theory-driven models by improving predictive accuracy while preserving interpretability. Emerging evidence also supports the integration of behavioural theories with ML frameworks. Azad et al. (2023) incorporated TPB constructs into multiple ML classifiers and found that models combining psychological variables with ML techniques achieved strong predictive performance while retaining explanatory value.

This study combines explanatory models with predictive models. TPB(Ajzen1991) has been extended to include three financial variables- saving motive, risk aversion and financial literacy. Primary data was collected from three Indian cities i.e. Bangalore, Chennai and Hyderabad. The study measured the impact of psychological variables (Attitude, subjective norms and perceived behavioural control) and financial variables (saving motive, risk aversion and financial literacy) on an individual's inclination to buy LI. Lastly, we studied whether saving motives mediates the link between financial literacy and purchase intention and whether risk aversion moderates this pathway. A predictive model using Random Forest and XGBoost was built.

Objectives of the Study

1. Examine how psychological constructs i.e. attitude, subjective norms, and perceived behavioural control impact life insurance purchase intention.
2. Assess the roles of financial constructs i.e. financial literacy, saving motive, and risk-aversion motive in influencing intention.
3. Study whether saving motive mediates the impact of financial literacy on purchase intention.
4. Evaluate whether risk aversion moderates the relationship between financial literacy and intention.
5. Use machine-learning models (Random Forest and XGBoost) to validate behavioural findings and compare predictive strengths.

Significance of the Study

The study has two major contributions to extant literature.

1. Extension of TPB by including financial variables to the research.
2. India's life insurance penetration remains low. Deeper insights into the factors shaping life insurance purchase decisions can enhance the effectiveness of insurance marketing efforts. This could result in better products, personalised communication and better targeting of life insurance products.

Literature Review

Despite raising income levels and better understanding of the benefits, life insurance sales have remained sluggish. This puzzling behaviour has prompted researchers to study what drives a person's intention to purchase life insurance.

Research studies have used Theory of planned behaviour to explain purchase intention of life insurance. TPB posits that intention is shaped by three psychological determinants **(Ajzen, 1991).** Recent studies have started taking a broader view. They argue that decisions about purchase of life insurance depends not only on psychological factors. Additional factors like what motivation to save, levels of financial literacy and risk orientation also impact the intention to purchase life insurance. (Hsu 2012; Niti et al., 2021). Also, there is increasing use of machine learning models in consumer finance research. (Chaudhuri & Singh, 2020). With all of this in mind, the following review traces what we know—and what we still do not know—about these drivers of insurance behaviour.

Most research studies use TPB to explain life insurance purchase behaviour. Among the TPB constructs, attitude exerted a statistically meaningful positive influence on purchase inclination. (Giri's 2018, Akiate et al. 2024). Akiate's study found that attitude has a higher impact on purchase intention compared to consumer's knowledge. Loke and Goh 2012 found that when consumers held a positive belief about life insurance, their intention to purchase life insurance increased considerably.

Subjective norms, which is the second TPB construct also has a strong effect on life insurance buying intention. This is an indication that social influence has a strong impact on life insurance purchase intention. (Sanjay & Tewari 2024, Gurang 2025, Pham et.al. 2024)

The third psychological variable is perceived behavioural control. Several research studies have shown that PBC has a major impact on buying inclination of LI. (Shafi et.al. 2010, Sanjay & Tewari 2025, Pratiwi 2014). When customers are confident that they understand and evaluate life insurance products, their PBC is high and so is their inclination to buy LI. The ability to understand is also linked to another variable i.e. financial literacy.

The next variable considered for the study is financial literacy. Research has shown that high levels of financial literacy result in better understanding, estimating and comparing benefits of life insurance products. Financial literacy also affects saving habits and long-term financial planning. (Gutter et.al. 2018, Agarwalla et.al 2015).

Other than risk mitigation, life insurance also ensures financial prudence, discipline and long-term savings. However, very few studies have studied saving motivation in the insurance context. (Thung 2012, Jain & Mandal 2022). Both the studies showed that saving motive has a strong impact on life insurance purchase intention. These studies didn't study the mediating effect of saving motive on financial literacy. This highlights an important lacuna in current scholarly research. Logically, financially literate consumers are more likely to save. No study has examined this linkage.

Traditional economic theory posits that risk aversion has a straightforward linkage with life insurance. The more risk averse a person, the more the likelihood of purchase of life insurance. This was confirmed by Browne and Kim 1993, who found that risk aversion had a strong effect on life insurance demand across countries. However recent studies found that the linage is a lot more complex. Nitti et.al 2021 found that risk aversion has a strong effect on the inclination to purchase life insurance. only when consumers financial literacy was moderate to high. So, risk aversion does not operate independently but depends on financial literacy. However, no study has formally tested this moderation effect and is a strong research gap.

Machine learning techniques are now being used by researchers to evaluate consumer behaviour. This has also been used in insurance research.ML algorithms like Random Forest and XG boost are now being used in consumer behaviour research. The reason is that these algorithms detect nonlinearities, and show which predictor is the most effective in predicting behaviour. This is often missed by traditional analysis like regression analysis (Chaudhuri & Singh, 2020). Studies which compared predictive accuracy, found that ML models yielded higher predictive accuracy. (Kwon and Kim 2022). While ML models are excellent in prediction, they cannot explain causal mechanisms. So, traditional behavioural models and machine learning can complement each other. Yet very few studies have used both techniques.

Literature review has shown the following: TPB constructs perform well in predicting intention. However, the accuracy varies across cultural and socio-economic settings. Financial literacy clearly impacts intention to purchase life insurance. But it is rarely used with an extended TPB framework that tests how it interacts with other variables. Saving motive shows theoretical promise but remains empirically underexplored, especially as a mediator. Risk aversion's influence is inconsistent and seems to depend on other factors, highlighting the need for formal moderation tests. And despite the growing popularity of machine-learning models, only a handful of studies have used them to validate behavioural findings, with none focusing specifically on Indian metropolitan consumers.

Research gaps: 1. No extended TPB framework which incorporates saving motive, risk aversion and financial literacy into one comprehensive model

2.No study combines explanatory behavioural modelling with ML based predictive modelling

3. We have no study on the mediating and moderating effects of financial motivators.

The present study aims to address these gaps. It proposes an extended TPB model which incorporates financial variables into the study. The study uses a motivator – saving motive, a psychological preference trait – risk aversion with a cognitive capability/knowledge variable- financial literacy. It also proposes use of ML models – Random Forest and XGBoost to validate the behavioural findings. This research adds to the growing body of literature by using a more comprehensive and a methodically blended approach to understand intention to purchase life insurance.

Methods

The research study uses a quantitative, descriptive, cross-sectional design. So, the data was gathered at one point. This gives us a snapshot of the variables of interest without trying to change or manipulate the environment. Since one of the aims of the study was to explore numerical associations; we adopted a quantitative approach to research design. Since the study didn't explore lived experiences, we didn't add qualitative methods to the study. Primary data was used for the study. A well-structured self-report questionnaire was used to collect data.

Data was collected across three cities of India i.e. Bengaluru, Chennai and Hyderabad. The choice of these cities was because of two reasons. These are the 4th,5th and 6th largest markets for life insurance in India. (IRDA 2024) Also, there is a large population of well-educated and professionally qualified population in these cities. This would be the ideal demographic for life insurance companies.

Data was collected over a two- month period (January – March 2025). Data was collected at universities and its affiliated centres, which helped us reach a diverse mix of respondents relatively quickly. The cross-sectional format felt appropriate for what we hoped to achieve. It allowed us to estimate the prevalence of key behavioural and financial variables and explore their interrelationships. This was achieved without intervening in people's decision environments. As methodological guidance consistently notes, the choice of design must follow the research question rather than the other way around (Guzmán-Muñoz et al., 2025).

Journal of Informatics Education and Research
ISSN: 1526-4726
Vol 6 Issue 1 (2026)

The target population included all students, faculty, and other staff at the university during the semester. This included students of part time programs for working professionals and short-term skill development courses. To create a proper sampling frame, we used official enrolment and employment records. Since the sampling frame was heterogenous, we opted for stratified random sampling method.

The students i.e. full time and part time were grouped as per their year of enrolment. Each of faculty and staff members were grouped as per the department. In each of the above groups, participants were chosen through simple random selection. This sampling technique ensured minimal selection bias. (Ahmed 2024)

Sample size: Sample size adequacy was assessed based on two principal factors. The first is that it should be able to detect meaningful results. This relates to statistical power. The other consideration is accuracy of results which is precision. An a priori power analysis for detecting medium-sized effects (e.g., Cohen's $d \approx 0.5$; $f^2 \approx 0.15$) at $\alpha = .05$ in regression-type models with multiple predictors indicates that samples in the low- to mid-300s typically achieve power of at least .80, especially when computed using established tools such as G*Power (Faul et al., 2007; Green, 1991; Suresh & Chandrashekara, 2012). Another popular tool to estimate sample size is Cochran's formula. This formula is used to ensure adequate precision in statistical estimations. Sample size as per Cochran's formula, estimated that a sample of about 385 cases is sufficient to estimate proportions with a ±5% margin of error at 95% confidence for large populations (Bartlett et al., 2001; Cochran, 1977). We also considered recent Monte Carlo work on structural equation models. These suggest that a sample size of 400 would provide stable parameter estimates. This would also provide an acceptable level of power of moderate complexity. (Wolf et al., 2013). So, we decided on a sample size of 400 completed questionnaires as a conservative and methodologically robust sample size. This sample size would exceed typical thresholds required for both statistical power and estimation precision.

We used a structured questionnaire to collect data. This was administered using Google Forms. The items were adapted from established and validated scales. We used an expert review to check content validity. We added a few items based on expert feedback. Additionally, a small pilot test was conducted ($n \approx 30$). Minor changes were made in the language and the layout based on the pilot studies results and feedback. Then final questionnaire was sent to 636 participants. We received 400 questionnaire that were suitable for analysis.

Data analysis was conducted using **JMP Pro 17 and SPSS**. After routine checks for missing data and outliers, we generated descriptive statistics to understand the sample's baseline characteristics. We examined normality and homoscedasticity assumptions and reported effect sizes and confidence intervals alongside p-values. Statistical significance was set at $p < .05$. Internal consistency was assessed using Cronbach's alpha, and all scales exceeded the commonly accepted 0.80 threshold, suggesting strong reliability.

As per our theoretical framework and study objectives, we proposed the following hypotheses:

- **H1:** Attitude has a significant positive effect on life insurance purchase intention.
- **H2:** Subjective norms have a significant positive effect on life insurance purchase intention.
- **H3:** Perceived behavioural control has a significant positive effect on life insurance purchase intention.
- **H4:** Financial literacy has a significant positive effect on life insurance purchase intention.
- **H5:** Saving motive has a significant positive effect on life insurance purchase intention.
- **H6:** Risk aversion has a significant positive effect on life insurance purchase intention.
- **H7:** Saving motive mediates the relationship between financial literacy and purchase intention.
- **H8:** Risk aversion moderates the relationship between saving motive and purchase intention.

We tested these hypotheses using Structural Equation Modelling (SEM) and followed this with predictive validation through machine-learning models.

Results Section

This section presents the comprehensive results of the analysis, including reliability and validity testing, structural model estimates, mediation and moderation analysis, and machine learning model performance.

Table 1. Reliability and Validity Statistics

Construct	Cronbach's Alpha	AVE	Composite Reliability
Attitude	0.81	0.62	0.84
Subjective Norms	0.78	0.59	0.82
Perceived Behavioural Control	0.80	0.61	0.83
Saving Motive	0.76	0.57	0.79
Risk Aversion	0.71	0.54	0.76
Financial Literacy	0.82	0.64	0.86

Table 1 presents the reliability and validity statistics. All constructs demonstrate satisfactory internal consistency and convergent validity. Cronbach's Alpha values exceed the 0.70 benchmark, Composite Reliability (CR) values surpass the required 0.70 standard, and AVE values are above the 0.50 threshold, indicating adequate convergent validity (Fornell & Larcker, 1981; Hair et al., 2019).

Table 2. Regression-Based Structural Model Results

Predictor	β Coefficient	Standard Error	t-value	p-value
Attitude	0.35	0.04	8.75	< .001
Subjective Norms	0.25	0.05	5.10	< .001
Perceived Behavioural Control	0.30	0.05	6.35	< .001
Saving Motive	0.20	0.05	4.20	< .001
Risk Aversion	0.15	0.06	1.98	0.048
Financial Literacy	0.25	0.05	5.25	< .001

Table 2 reports the structural model estimates. All hypothesised paths are significant at $p < .001$ except risk aversion which is marginally significant at $p = .048$. Attitude, subjective norms, perceived behavioural control, financial literacy, and saving motive exhibit strong positive effects, while risk aversion plays a relatively weaker role. These effects are theoretically aligned with TPB foundations and financial behaviour literature (Ajzen, 1991; Armitage & Conner, 2001).

Table 3. Mediation Analysis (Financial Literacy → Saving Motive → Purchase Intention)

Path	Effect	Standard Error	p-value
Financial Literacy → Saving Motive (a)	0.41	0.06	< .001
Saving Motive → Intention (b)	0.20	0.05	< .001

Indirect Effect (ab)	0.082	0.02	< .01

Table 3 shows mediation analysis, confirming that saving motive significantly mediates the effect of financial literacy on purchase intention. Although mediation is statistically significant, future work should consider bootstrapped confidence intervals to reinforce inferential strength (Hayes, 2018).

Table 4. Moderation Analysis (Risk Aversion × Saving Motive)

Predictor	β Coefficient	t-value	p-value
Saving Motive	0.20	4.20	< .001
Risk Aversion	0.15	1.98	0.048
Interaction Term	0.03	0.82	0.41

Table 4 presents moderation analysis, indicating that the interaction term is insignificant. This suggests that risk aversion does not moderate the saving motive–intention relationship, consistent with financial-behaviour studies where risk aversion shows weak moderating influence.

Table 5. Machine Learning Regression Performance

Model	RMSE	MAE	R^2
XGBoost	**0.566**	**0.452**	**0.745**
Random Forest	0.523	0.401	0.781

Table 6. Machine Learning Classification Performance

Model	Accuracy	Precision	Recall	F1-score	AUC
XGBoost	0.758	0.782	0.717	0.748	0.840
Random Forest	0.758	0.772	0.733	0.752	0.857

Tables 5 and 6 display machine learning model performance, where Random Forest and XGBoost exhibit high predictive accuracy, exceeding typical SEM explanatory benchmarks. Table 7 shows feature importance rankings, which closely align with SEM predictors.

Table 7. Feature Importance (Random Forest Model)

Feature	Importance Score
Attitude	0.262
Subjective Norms	0.189
Perceived Behavioural Control	0.170
Saving Motive	0.120
Financial Literacy	0.142
Risk Aversion	0.117

Table 8. Discriminant Validity – Fornell–Larcker Criterion

Note: Diagonal values (bold) are AVE. Off-diagonal values are inter-construct correlations. Discriminant validity is confirmed when diagonal values exceed corresponding correlations (Fornell & Larcker, 1981).

Construct	Attitude	Subjective Norms	Perceived Behavioral Control	Saving Motive	Risk Aversion	Financial Literacy
Attitude	0.787	0.1468	0.0033	0.2654	0.0004	0.0703
Subjective Norms	0.3832	0.768	0.0032	0.033	0.019	0.014
Perceived Behavioral Control	-0.0573	0.0564	0.781	0	0.2409	0.0006
Saving Motive	0.5152	0.1816	-0.0051	0.755	0.0121	0.2686
Risk Aversion	0.0206	0.1378	0.4908	0.1099	0.735	0.0093
Financial Literacy	0.2652	0.1182	0.0235	0.5182	0.0963	0.800

For all constructs, the square root of AVE (diagonal values ≈ 0.735–0.800) is greater than the corresponding inter-construct correlations. Thus, discriminant validity is satisfied in accordance with Fornell and Larcker (1981) and Rönkkö and Cho (2022).

Table 9. Discriminant Validity – HTMT Ratios

Note: HTMT < 0.85 indicates conservative discriminant validity; < 0.90 indicates acceptable discriminant validity (Henseler et al., 2015).

Construct	Attitude	Subjective Norms	Perceived Behavioral Control	Saving Motive	Risk Aversion	Financial Literacy
Attitude	1.000					
Subjective Norms	0.714	1.000				
Perceived Behavioural Control	0.573	0.754	1.000			
Saving Motive	0.618	0.39	0.584	1.000		
Risk Aversion	0.609	0.763	0.817	0.595	1.000	
Financial Literacy	0.520	0.663	0.625	0.699	0.692	1.000

Discriminant validity was confirmed through both Fornell–Larcker Criterion and HTMT analysis (Table 8 and Table 9), confirming constructs are empirically distinct. As R^2 values of 0.50 and above are typically interpreted as indicating at least moderate, and often substantial, explanatory power in social science models (Chin, 1998; Hair et al., 2017, 2021), the $R^2 = 0.64$ obtained for life-insurance purchase intention can be considered more than acceptable.

All HTMT values fall below the conservative 0.85 threshold; the highest HTMT value is **0.817**, which remains within acceptable discriminant validity limits. Therefore, discriminant validity is strongly supported (Henseler et al., 2015; Roemer & Henseler, 2021).

Discussion

Table 1 shows that the reliability and validity statistics. As all the values are higher than the minimum prescribed thresholds, it can be inferred that the measuring instrument has exhibited required validity and reliability.

Table 2 shows the results of regression analysis of SEM results. All hypothesised paths of the SEM model are significant at $p < .001$ except risk aversion which is marginally significant at $p = .048$. Attitude, subjective norms, perceived behavioural control, financial literacy, and saving motive exhibit strong positive effects, while risk aversion plays a relatively weaker role. These effects are theoretically aligned with TPB foundations and financial behaviour literature (Ajzen, 1991; Armitage & Conner, 2001). This is also aligned with other research studies on purchase intention of life insurance. (Nomi & Sabbir 2020, Omar et.al 2007, Reshma & Shacheendran (2023)).

Based on SEM and supporting machine-learning results (Table 2 and table 5) the hypotheses were evaluated as follows:

H1: Accepted the hypothesis — Attitude significantly predicts intention.
H2: Accepted the hypothesis — Subjective norms significantly influence intention.
H3: Accepted the hypothesis — Perceived behavioural control significantly predicts intention.
H4: Accepted the hypothesis — Financial literacy has a strong significant effect.
H5: Accepted the hypothesis — Saving motive significantly predicts intention.
H6: Hypothesis is accepted as the effect is significant. But the effect is lower than the other variables — Risk aversion has a positive but marginal effect.
H7: Accepted the hypothesis— Saving motive mediates the effect of financial literacy.
H8: Hypothesis is rejected — Interaction effect was non-significant; risk aversion does not moderate the relationship.

This shows a strong SEM model and highlights mediation effects.

Results of SEM, along with machine learning results (Table 2 &5) show that the three core variables of TPB i.e. Attitude, Subjective norms and PBC all showed a significant and positive effect on the purchase intention of LI. The results reaffirm decades of TPB research. (Ajzen, 1991; Armitage & Conner, 2001). Attitude was the strongest predictor. This is in line with earlier research studies. (Nomi & Sabbir 2020, Omar et.al 2007, Reshma & Shacheendran 2023, Sanjay & Tewari 2025). This demonstrates that if a consumer is favourably inclined to towards a product, the higher would be the intention to purchase it. Subjective norms and PBC also have a significant impact. This shows that social influences and perceived ease of purchase also significantly affect purchase intention. (Bansal & Taylor, 2002; George, 2004).

We had extended TPB by adding three additional predictors. Saving motive had a strong and positive impact on purchase intention of life insurance. The findings are consistent with other studies. (Mahdzan & Victorian, 2013; Truong & McColl, 2011). Life insurance ensures long term and disciplined saving. This is a powerful motivator to purchase life insurance.

The second extended predictor was financial literacy. Results show that financial literacy has a strong and positive effect on the intention to purchase life insurance. This is aligned to findings of earlier studies. (Zakaria et al. (2016), Djoni and Rahardjo (2021). However, other studies show that financial literacy does not have a significant direct effect but has a mediating role especially on attitude. (Gurang et.al 2025, Sanjay & Tewari 2025). So, the effect of financial literacy is more complex. The role it assumes depends on the context i.e. age, education levels, culture etc of the sample.

The third extended parameter was risk aversion. It has a relatively weaker but significant effect on the intention to purchase life insurance. This is still in line with yet still in line with the classical economic view of life insurance. This view is that life insurance is primarily an instrument for financial risk mitigation. So higher the risk aversion the higher the inclination of that individual to purchase life insurance. (Browne & Kim, 1993; Outreville, 2015). Results of this study show that saving motivation is higher than risk aversion motivation. This could be a country specific result as LI is primarily sold in India as a saving and an income tax reduction instrument.

One of the novel and interesting findings of this study is the results of the mediation effects as shown in table 3. Saving motive mediates the effect of financial literacy. It acts as a bridge between financial literacy and purchase intention. This is also intuitive as individuals whose financial literacy is high recognise the importance of disciplined long-term savings. This makes life insurance an attractive proposition. So, saving motives acts as a psychological channel through which financial literacy strongly influences buying intention of life insurance. This aligns with earlier work showing that saving motives strongly influence insurance behaviour (Sanjay & Tewari, 2024).

The results of the moderation analysis are as shown in table 4. The results shows that risk aversion neither strengthened nor weakened the link between saving motive and intention. As can be seen in table 4, the interaction term is not significant. This suggests that even individuals with a high-risk aversion did not respond differently to their saving motives. Other studies have also shown that risk aversion shows weak or unstable moderating effects in financial models. (Donkers et al., 2001; Hallahan et al., 2004). Our study is aligned with other studies that shows that risk aversion influences behaviour directly (Sitkin & Pablo, 1992; Weber et al., 2002). However, it does not significantly alter other motivations like saving motives.

Let’s now look at a comparison with extant literature. The strong effect of the three variables of TPB is in line with earlier studies which showed that positive attitude, strong subjective norms and high PBC score enhance purchase intentions. (Chen & Tung, 2014; Kang et al., 2013). The significance of saving motive and financial literacy is also in line with extant literature. This study findings diverge from Sanjay & Tewari 2024 and Mabula & Ping, 2023 who reported no significant effect of financial literacy on intention. This discrepancy could be due to contextual differences or could point to more a more complex relationship.

Risk aversion shows interesting results in comparison with earlier scholarship. Studies based on economic theory (Yaari1965) and some of the later studies (Barsky et al., 1997; Outreville, 2015) showed that higher the risk aversion, higher the intention to purchase life insurance. However, other studies show that the effect of risk aversion on intention is modest. This is especially true when you take into account other factors like saving motive and financial literacy .(Giné et al., 2008). This is further reinforced by the lack of moderation effect. This is in line with earlier studies that found that risk aversion rarely produces strong effect in the context of intention to purchase financial products. (Hallahan et al., 2004; Donkers et al., 2001).

One of the novel features of this study in the use of machine -learning models. The results are shown in tables 5,6 &7. Both the ML models i.e. Random forest and XGBoost significantly outperformed SEM in predicting purchase intention. Random Forest achieved an R^2 of 0.781 and XGBoost 0.745, well above the SEM’s ~0.63. This is in line with other studies which show that ensemble models like Random forests and XGBoost often capture nonlinearities and complex interactions and so are more effective than linear models. (Kuhn & Johnson, 2013; Shmueli et al., 2016). The classification performance was similar with both models performing well with Random Forest performing better. Feature importance rankings largely mirrored the SEM results. Attitude had the highest ranking followed by subjective norms, PBC and the other financial variables. This is in line with other studies which show that attitude is the strongest predictor. (Armitage & Conner, 2001).

The performance parameters are as shown in table 4. These further show the robustness of the model. Both Random Forest and XGBoost demonstrated strong predictive capability. Though both the models performed very well, Random Forest marginally outperformed XGBoost across most indicators. The superior performance of both the ML models over SEM demonstrates their superior ability to capture nonlinear relationships and complex interaction patterns which are often overlooked by traditional statistical models.

Table 7 shows the feature importance hierarchy. It closely mirrors the results of the SEM. Attitude is the strongest predictor followed by subjective norms, PBC. Saving motive and financial literacy also show significant predictive capability. The study shows that in this context there is a convergence between the theory driven explanatory models of SEM and the data driven approach of the ML models. This demonstrates compelling evidence that the behavioural mechanisms identified in the study meaningfully translate into predictive accuracy. This reinforces the credibility and applicability of the model in the real world in the context of life insurance.

The strong performance of the ML models indicates that there are non-linearities or certain interactions which is not detected by traditional SEM. While SEM reveals theoretical pathways, ML clarifies predictive hierarchy and pattern strength. So, both the models complement each other and show a more comprehensive picture than that using one of the methods.

These results have extended the TPB framework by showing how financial motivation constructs enrich and offer new insights to the process of formation of purchase intention. The mediation effect i.e. financial literacy → saving motive → intention shows a meaningful psychological mechanism. It aligns with study which shows that better financial knowledge fosters saving oriented thinking. (Lusardi & Mitchell, 2014; Xiao & Porto, 2017). This suggests that further studies using TPB in a financial context should include domain specific variables like saving motives and financial literacy. These studies have recommended expanded TPB models in research studies in financial domain. (East, 1993; Yazdanpanah & Forouzani, 2015).

The findings point toward a richer, more layered understanding of how individuals form intentions to purchase life insurance. Attitude remains central, but financial literacy and saving motives provide important additional texture, while machine-learning results underscore the complexity of behavioural prediction in this domain.

Conclusion

This study set out to address five key research objectives: to test the influence of the three core TPB constructs on life insurance purchase intention; to assess the role of financial literacy, saving motive, and risk aversion; to examine mediation and moderation pathways; and to validate behavioural findings through machine-learning models. The results of the study show that all the objectives have been fulfilled. The three psychological constructs of TPB significantly affect purchase intention. Among these constructs, attitude is the strongest predictor. This confirms that purchase intention of life insurance is strongly affected by psychological evaluations, social pressures and perceived capability. Thus, objective 1 has been achieved.

Financial literacy and saving motive significantly shape intention, while risk aversion exerts a weaker effect, addressing Objective 2. The third objective was about mediation effects. The mediation analysis shows that saving motive acts as a psychological channel through which financial literacy strengthens intention, fulfilling the objective 3. Mediation analysis showed that risk aversion does not moderate the pathway. This suggests that its influence is largely direct rather than interactive (Objective 4). Finally, Random Forest and XGBoost demonstrated high predictive power, validating behavioural results and revealing nonlinear dynamics, thereby achieving Objective 5.

These findings have direct relevance for addressing India's persistent low insurance penetration. First, strengthening favourable attitudes toward insurance must become a central communication priority. Public awareness campaigns, IRDAI outreach programmes, and insurer messaging should move beyond product description to emotionally anchor life insurance as a responsible, future-protective decision. Second, perceived behavioural control suggests that simplifying policies, reducing documentation burden, enhancing digital transparency, and improving agent credibility can materially enhance purchase intention. Third, the strong role of financial literacy and saving motive indicates that financial education cannot merely be generic; it must explicitly connect literacy to disciplined long-term saving behaviour and demonstrate how life insurance aligns with life-cycle financial goals. Embedding insurance awareness in financial-literacy curricula, workplace financial programmes, and national education initiatives can therefore yield meaningful behavioural movement. Fourth, the weaker role of risk aversion indicates that urban consumers increasingly perceive life insurance as a savings-cum-investment product. Policymakers and insurers must acknowledge this reality by designing clearer hybrid products, ensuring transparent returns communication, and framing insurance not only as risk protection but as structured long-term financial planning.

Finally, the convergence of SEM and machine-learning results indicates that policy interventions can be data-driven and behaviourally grounded rather than assumption-led. By combining behavioural theory with predictive analytics, policymakers, insurers, and regulators can better segment consumers, design targeted nudges, and close India's significant protection gap. This study shows that improving life insurance penetration requires a multi-dimensional strategy—strengthening attitudes, empowering consumers, deepening financial capability, and aligning policy design with behavioural realities rather than purely economic assumptions.

References

1. Agarwalla, S. K., Barua, S., Jacob, J., & Varma, J. R. (2015). Financial literacy among working young in urban India. *Economic and Political Weekly*, 50(48), 51–59.
2. Ahmed, S. (2024). Stratified sampling methods in social science research. Journal of Applied Statistics, 51(3), 455–468.
3. Akiate, DBA et.al. THE EFFECT OF CONSUMER KNOWLEDGE AND ATTITUDE ABOUT LIFE INSURANCE ON THE PURCHASE INTENTION OF ACCOUNTANTS WITHIN PAMPANGA. International Journal of Engineering Applied Sciences and Technology. 09. 26-35. 10.33564/IJEAST.2024.v09i01.003.
4. Ajzen, I. (1991). The theory of planned behavior. Organizational Behavior and Human Decision Processes, 50(2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T
5. Armitage, C. J., & Conner, M. (2001). Efficacy of the theory of planned behaviour: A meta-analytic review. British Journal of Social Psychology, 40(4), 471–499. https://doi.org/10.1348/014466601164939
6. Azad, Md Shawmoon, et al. "Predictive modeling of consumer purchase behavior on social media: Integrating theory of planned behavior and machine learning for actionable insights." *Plos one* 18.12 (2023): e0296336.
7. Azzone, Michele, et al. "A machine learning model for lapse prediction in life insurance contracts." *Expert Systems with Applications* 191 (2022): 116261.
8. Bansal, H. S., & Taylor, S. F. (2002). Investigating interactive effects in the theory of planned behavior. Psychology & Marketing, 19(9), 693–718.
9. Bartlett, J. E., Kotrlik, J. W., & Higgins, C. C. (2001). Organizational research: Determining appropriate sample size in survey research. Information Technology, Learning, and Performance Journal, 19(1), 43–50.
10. Barsky, R. B., Juster, F. T., Kimball, M. S., & Shapiro, M. D. (1997). Preference parameters and behavioral heterogeneity: An experimental approach in the health and retirement study. *The quarterly journal of economics*, *112*(2), 537-579.
11. Beck, T., & Webb, I. (2003). Economic, demographic, and institutional determinants of life insurance consumption across countries. World Bank Economic Review, 17(1), 51–88.
12. Bhat, A., & Jain, R. (2013). Life insurance purchase behaviour in India. Journal of Insurance Studies, 5(2), 34–49.
13. Bhatia, A., Jain, S., & Mehta, P. (2021). Predicting insurance purchase using machine learning. Journal of Financial Data Science, 3(4), 66–81.
14. Browne, M. J., & Kim, K. (1993). An international analysis of life insurance demand. Journal of Risk and Insurance, 60(4), 616–634.
15. Chaudhuri, S., & Singh, R. (2020). Machine learning applications in consumer finance. Journal of Financial Innovation, 6(2), 1–15.
16. Chen, M. F., & Tung, P. J. (2014). Developing an extended theory of planned behavior model to predict consumers' intention to visit green hotels. *International journal of hospitality management*, *36*, 221-230.
17. Chin, W. W. (1998). The partial least squares approach to structural equation modeling. In G. A. Marcoulides (Ed.), Modern methods for business research (pp. 295–336). Lawrence Erlbaum.
18. Cochran, W. G. (1977). Sampling techniques (3rd ed.). Wiley.
19. Dash, M., & Mahapatra, S. (2014). Determinants of life insurance demand in India. Journal of Risk Finance, 15(2), 145–160.
20. Djoni, K., & Rahardjo, S. (2021). *Analysis of financial literacy, income and premium on the purchase intention of life insurance*. Budapest International Research and Critics Institute Journal, 4(3), 3216–3226. ([Bircu Journal][3])
21. Donkers, B., Melenberg, B., & Van Soest, A. (2001). Estimating risk attitudes. Journal of Economic Psychology, 22(4), 1–20.
22. East, R. (1993). Investment decisions and the theory of planned behaviour. *Journal of economic psychology*, *14*(2), 337-375.
23. Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3. Behavior Research Methods, 39(2), 175–191.
24. Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables. Journal of Marketing Research, 18(1), 39–50.

25. George, J. F. (2004). The theory of planned behavior and Internet purchasing. *Internet research*, *14*(3), 198-212.
26. Ghafar, Fabu Ayub A., Norhana Abd Rahim, and Yumn Suhaylah Yusoff. "Predicting Customer Behaviour on Buying Life Insurance using Machine Learning." *Journal of Advanced Research in Marketing and Management* 2.1 (2025): 39-50.
27. Giné, X., Townsend, R., & Vickery, J. (2008). Patterns of risk sharing. Journal of Development Economics, 85(1–2), 234–256.
28. Giri, A., & Chatterjee, S. (2021). Life insurance persistence in India. Journal of Financial Services Marketing, 26(4), 243–257.
29. Green, S. B. (1991). How many subjects does it take to do a regression analysis? Multivariate Behavioral Research, 26(3), 499–510.
30. Gramegna, Alex, and Paolo Giudici. "Why to buy insurance? An explainable artificial intelligence approach." *Risks* 8.4 (2020): 137.
31. Gurung, G., Ojha, A., Kharel, S., & Simkhada, B. (2025). Drivers of consumer purchase intention toward life insurance. Nepalese Journal of Insurance and Social Security, 8(1), 43–53.
32. Gutter, M. S., Copur, Z., & Fontes, A. (2018). Financial capabilities and financial behaviors of low-income households. *Journal of Consumer Affairs*, 52(1), 144–165.
33. Guzmán-Muñoz, E., Mendez-Rebolledo, G., Concha-Cisternas, Y., Alarcón-Rivera, M., & Faúndez-Casanova, C. (2025). Diseños de investigación cuantitativa en ciencias de la actividad física y la salud. *Revista Ciencias de la Actividad Física UCM*, *26*(2), 63-85.
34. Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2017). A primer on partial least squares structural equation modeling (PLS-SEM) (2nd ed.). Sage.
35. Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European business review*, *31*(1), 2-24.
36. Hair, J. F., Jr., Hult, G. T. M., Ringle, C. M., Sarstedt, M., & Thiele, K. O. (2021). Evaluation of the structural model. In *Partial least squares structural equation modeling (PLS-SEM) using R* (pp. 123–142). Springer. https://doi.org/10.1007/978-3-030-80519-7_6
37. Hallahan, T. A., Faff, R. W., & McKenzie, M. D. (2004). An empirical investigation of personal financial risk tolerance. Financial Services Review, 13(1), 57–78.
38. Hayes, A. F. (2018). Introduction to mediation, moderation, and conditional process analysis (2nd ed.). Guilford Press.
39. Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity. Journal of the Academy of Marketing Science, 43(1), 115–135.
40. Hsu, J., & Willis, R. (2013). Dementia risk and financial decision making. *Journal of the Economics of Ageing*, 1–2, 27–38.
41. .Insurance Regulatory and Development Authority of India (IRDAI). (2025). *Annual Report 2024-2025*. IRDAI
42. Jain, R., & Mandal, P. (2022). Savings behaviour and insurance adoption among Indian millennials. *Journal of Financial Planning and Analysis*, 5(2), 112–128.
43. Kahneman, D. (2011). Thinking, fast and slow. Farrar, Straus and Giroux.
44. Kakar, Preeti, and Rajesh Shukla. "The determinants of demand for life insurance in an emerging economy—India." *Margin: The Journal of Applied Economic Research* 4.1 (2010): 49-77.
45. Kang, E., & Yoon, S. (2024). Does technology-based self-service (TBSS) help tourists to revisit destinations? Validation of extended theory of planned behavior. *Global Business & Finance Review (GBFR)*, *29*(7), 48-62.
46. Kuhn, M., & Johnson, K. (2013). Applied predictive modeling. Springer.
47. Kwon, H., & Kim, S. (2022). Comparative analysis of machine-learning and behavioural models in predicting insurance purchase. *Insurance Markets and Companies*, 13(1), 1–12.
48. Liedtke, P. M. (2007). What does insurance contribute to economic growth? Geneva Papers on Risk and Insurance, 32(2), 221–239.
49. Loke, Y. J., & Goh, Y. Y. (2013). Purchase Decision of Life Insurance Policies among Malaysians. International Journalof Social Science and Humanity, 2(5), 415–420. https://doi.org/10.7763/IJSSH.2012.V2.137
50. Lusardi, A., & Mitchell, O. S. (2014). The economic importance of financial literacy. Journal of Economic Literature, 52(1), 5–44.

51. Mabula, J. B., Dong Ping, H., & James, M. (2023). The impact of African Firms' utilization of financial and technology resource on innovation: a simple mediation. *Sage Open*, *13*(1), 21582440231153037.
52. Mahdzan, N. S., & Victorian, S. M. P. (2013). Determinants of life insurance demand. Asian Social Science, 9(16), 51–59.
53. Nomi, Marzia, and Md Mahiuddin Sabbir. "Investigating the factors of consumers' purchase intention towards life insurance in Bangladesh: An application of the theory of reasoned action." *Asian Academy of Management Journal* 25.2 (2020).
54. Niti, P., Sharma, R., & Chawla, D. (2021). Risk attitudes, financial literacy and insurance decisions in India. *International Journal of Consumer Studies*, 45(6), 1231–1245.
55. Ejye Omar, O., & Owusu-Frimpong, N. (2007). Life insurance in Nigeria: An application of the theory of reasoned action to consumers' attitudes and purchase intention. *The Service Industries Journal*, *27*(7), 963-976.
56. Outreville, J. F. (2015). Risk aversion and insurance demand. Journal of Insurance Issues, 38(1), 1–20.
57. Pauly, M. V., Kunreuther, H., & Hirth, R. (2003). Guaranteed renewability in insurance. Journal of Risk and Insurance, 70(1), 1–20.
58. Pham, T. T. T., To, T. H., & Le, T. H. T. (2023). A study on factors affecting intentions to purchase life insurance through bancassurance channel in Vietnam. *Journal of Economics, Finance and Management Studies, 6*(10), 5050–5056. (ResearchGate)
59. Pratiwi, N. A. (2014). *Analisis niat beli asuransi jiwa pada mahasiswa menggunakan teori rencana perilaku* [Unpublished manuscript]. Institut Pertanian Bogor. ([Journal IPB][2])
60. Reshma, K. R., & Shakheendran, V. (2023). Planned behaviour in purchasing postal life insurance: An empirical study. *Bimaquest, 23*(2), 1–23. (bimaquest.niapune.org.in)
61. Rönkkö, M., & Cho, E. (2022). An updated guideline for assessing discriminant validity. *Organizational Research Methods, 25*(1), 6–34. https://doi.org/10.1177/1094428120968614
62. Roemer, E., & Henseler, J. (2021). HTMT2—An improved criterion for assessing discriminant validity in structural equation modeling. *Industrial Management & Data Systems, 121*(12), 2637–2650. https://doi.org/10.1108/IMDS-02-2021-0082
63. Sanjay, K., & Tewari, S. (2024). Analysing the determinants of life insurance purchase intention. Srusti Management Review, 17(2), 192–207.
64. Sanjay, K., & Tewari, S. (2025). Financial motivators and behavioural intention. Srusti Management Review, 18(1), 33–52.
65. SBI Life Insurance & Deloitte India. (2023, September 11). *Financial Immunity Study 3.0: Demystifying the Indian Consumers' Illusions* [Report]. SBI Life Insurance. https://www.sbilife.co.in/financialimmunity
66. Shmueli, G., Ray, S., Velasquez Estrada, J. M., & Chatla, S. (2016). The elephant in the room: Predictive performance of PLS models. *Journal of Business Research, 69*(10), 4552–4564. https://doi.org/10.1016/j.jbusres.2016.03.049
67. Shamsuddin, A., Rahman, M., & Ali, S. (2023). Gradient boosting in insurance analytics. Expert Systems with Applications, 212, 118708.
68. Shafi, M., Afridi, S., & Naeem, M. (2021). Determinants of life insurance demand in urban India. *Asian Journal of Business Research*, 11(2), 45–62.
69. Shmueli, G., Ray, S., Velasquez Estrada, J. M., & Chatla, S. (2016). The elephant in the room. Journal of Business Research, 69(10), 4552–4564.
70. Sitkin, S. B., & Pablo, A. L. (1992). Reconceptualizing the determinants of risk behavior. *Academy of management review*, *17*(1), 9-38.
71. Suresh, K. P., & Chandrashekara, S. (2012). Sample size estimation. Journal of Human Reproductive Sciences, 5(1), 7–13.
72. Thaler, R. H., & Sunstein, C. R. (2008). Nudge. Yale University Press.
73. Thung, C. M., et al. (2012). Determinants of saving behaviour among Malaysian young adults. *Journal of Economics and Behavioral Studies*, 4(1), 33–41.
74. Truong, Y., & McColl, R. (2011). Intrinsic motivations, self-esteem, and luxury goods consumption. *Journal of retailing and consumer services*, *18*(6), 555-561.

75. Weber, E. U., Blais, A. R., & Betz, N. E. (2002). A domain-specific risk-attitude scale: Measuring risk perceptions and risk behaviors. *Journal of behavioral decision making*, *15*(4), 263-290.
76. Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013). Sample size requirements for SEM. Educational and Psychological Measurement, 73(6), 913–934.
77. Xiao, J. J., & Porto, N. (2017). Financial education and financial satisfaction: Financial literacy, behavior, and capability as mediators. *International Journal of Bank Marketing*, *35*(5), 805-817.
78. Yaari, M. E. (1965). Uncertain lifetime and life insurance. Review of Economic Studies, 32(2), 137–150.
79. Yazdanpanah, M., & Forouzani, M. (2015). Application of the Theory of Planned Behaviour to predict Iranian students' intention to purchase organic food. *Journal of Cleaner Production*, *107*, 342-352.
80. Zakaria, Z., Azmi, N. M., Nik Hassan, N. F. H., Salleh, W. A., Tajuddin, M. T. H. M., Sallem, N. R. M., & Noor, J. M. M. (2016). The intention to purchase life insurance: A case study of staff in public universities. *Procedia Economics and Finance, 37*, 358–365. [https://doi.org/10.1016/S2212-5671(16)30137-X](https://doi.org/10.1016/S2212-5671%2816%2930137-X) ([ScienceDirect][1])