FIN 450 – Finance and Artificial Intelligence Exam 2 Study Guide

Use this as your master review for the exam. Focus on the Key Takeaways, Common Pitfalls, and Exam Cram Sheet when time is short.

Topic 9 – Risk, Return, and Performance Metrics

9.1 Key Concepts & Definitions

- Holding Period Return (HPR) – total return over a specific period (income + price change) divided by initial value.
- Equivalent Annual Return (EAR) – annualized return that would give the same multi-period result.
- Arithmetic Mean Return – simple average of periodic returns.
- Geometric Mean Return / CAGR – compounded average return over time.
- Excess Return – return minus a benchmark (risk-free or index).
- Variance / Standard Deviation (σ) – dispersion of returns around their mean.
- Covariance – how two returns move together in level and direction.
- Correlation (ρ) – scaled covariance; strength of linear co-movement (-1 to $+1$).
- R-Squared (R^2) – fraction of variance explained by a model (e.g., CAPM).
- Beta (β) – sensitivity of an asset to the market factor.
- Max Drawdown (MDD) – largest peak-to-trough loss.
- Value at Risk (VaR) – threshold loss over horizon at given confidence (e.g., 99%/10-day).
- Semi-Deviation / Downside Deviation – volatility of returns below a target.
- Sharpe Ratio – excess return per unit of total volatility.
- Sortino Ratio – excess return per unit of downside volatility.

9.2 Core Principles

- Financial analysis is done in returns, not dollar changes.
- Use arithmetic mean for one-period expectations; geometric mean for multi-year wealth.
- Risk has many faces: volatility, tail risk, drawdown, downside risk.

- Risk-adjusted measures (Sharpe/Sortino) compare reward per unit of risk.
- Always annualize consistently (returns, σ, Sharpe).

9.3 Technical Details

- Arithmetic vs Geometric Returns: arithmetic is the simple average; geometric is the compounded average.
- When volatility is positive, arithmetic mean is greater than geometric mean.
- Annualization: returns via compounding, volatility via square root of time.
- Sample variance: $s^2 = (1/(N-1)) \Sigma (r_t - \bar{r})^2$.
- Covariance: $Cov(X,Y) = (1/(N-1)) \Sigma (x_t - \bar{x})(y_t - \bar{y})$.
- Correlation: $\rho = Cov(X,Y) / (\sigma_X \sigma_Y)$.
- Beta (β) from regression of asset on market: $\beta = Cov(r_i,r_m) / \sigma_m^2$.
- Sharpe Ratio: $S = (\bar{r}_p - r_f) / \sigma_p$.
- Sortino Ratio uses downside deviation instead of total σ.
- Parametric VaR (normal): use mean and σ to estimate a loss quantile.

9.4 Real-World Applications

- Portfolio performance reports for mutual funds and hedge funds.
- Risk reports: σ, β, MDD, VaR for regulators and clients.
- Comparing managers using Sharpe, Sortino, and tracking error.
- Lab context: using AI/Excel to compute returns, risk, performance and checking calculations by hand.

9.5 Important Distinctions

- Arithmetic vs Geometric Mean: expectation vs long-run growth.
- Standard Deviation vs Max Drawdown: day-to-day variability vs worst episode.
- Sharpe vs Sortino: total volatility vs downside-only volatility.
- Covariance vs Correlation: raw co-movement vs scaled, unit-free measure.

9.6 Key Takeaways – Topic 9

- Know which mean (arithmetic or geometric) you are using and why.
- Volatility is not the whole story: drawdown and downside risk matter.
- Sharpe and Sortino are central for comparing strategies.
- Always check AI-produced numbers in Excel or by hand.

9.7 Common Pitfalls – Topic 9

- Mixing non-annualized with annualized numbers.
- Using arithmetic mean for multi-year growth questions.

- Forgetting to subtract the risk-free rate when computing excess returns.
- Treating correlation as stable during crises.

Topic 10 – Risk Analysis (Sensitivity, Scenario, Monte Carlo)

10.1 Key Concepts & Definitions

- Sensitivity Analysis – vary one (or two) inputs, see effect on output.
- Scenario Analysis – evaluate coherent stories where multiple inputs change together.
- Monte Carlo (MC) Simulation – repeatedly sample from distributions for uncertain inputs to estimate the distribution of outcomes.
- Elasticity – percentage response of output to a percentage change in input.
- Tornado Chart – bar chart ranking variables by impact on NPV.
- Spider Chart – NPV vs percentage change in each input; slopes show sensitivity.

10.2 Core Principles

- Risk analysis answers systematic what-if questions.
- Sensitivity analysis is local and one-driver-at-a-time.
- Scenario analysis uses narrative bundles of drivers (base/upside/downside).
- Monte Carlo simulation gives a full distribution of NPV, not just a few points.
- AI can enhance simulation speed, realism, and interpretation.

10.3 Technical Details

- Sensitivity analysis: start from base case and vary one driver (e.g., price ±10%).
- One-way and two-way data tables in Excel are common for sensitivity analysis.
- Scenario analysis: define base, upside, and downside cases with coherent assumptions.
- Monte Carlo simulation steps: define distributions, simulate inputs, compute outputs, summarize distribution.
- Monte Carlo outputs: mean NPV, σ, percentiles (P10, P50, P90), VaR, CVaR.

10.4 Summary Table – Risk Analysis Methods

Sensitivity Analysis: varies one input at a time; simple and transparent but ignores joint probabilities.

Scenario Analysis: varies multiple inputs in coherent stories; intuitive but subjective and may miss tail outcomes.

Monte Carlo Simulation: varies all uncertain inputs via distributions; captures probabilities and tails but is data- and model-intensive.

10.5 Real-World Applications

- Capital budgeting risk analysis in energy, real estate, pharma, and tech.
- Regulatory stress testing for banks and insurance companies.
- Portfolio risk under FX or volatility shocks.
- Using AI to speed Monte Carlo simulation and identify key drivers.

10.6 Key Takeaways – Topic 10

- Know when to use sensitivity vs scenario vs Monte Carlo analysis.
- Be able to explain the pros and cons of each method.
- Understand how to interpret percentiles and VaR/CVaR from Monte Carlo output.

10.7 Common Pitfalls – Topic 10

- Treating sensitivity analysis results as probabilities.
- Choosing unrealistic or inconsistent scenario assumptions.
- Ignoring dependencies/correlations in Monte Carlo models.
- Over-trusting AI/Monte Carlo outputs without sanity checks.

Topic 11 – NLP in Finance

11.1 Key Concepts & Definitions

- Natural Language Processing (NLP) – methods to turn text (and speech) into structured signals.
- TF-IDF – term frequency–inverse document frequency representation of text.
- Sentiment Analysis – classifying tone (positive/neutral/negative).
- FinBERT – finance-tuned transformer for sentiment and classification.
- F1 Score – harmonic mean of precision and recall; measures classification quality.
- Named Entity Recognition (NER) – tagging entities like companies, people, places, securities.
- RAG (Retrieval-Augmented Generation) – LLM plus document retrieval.

- Automatic Speech Recognition (ASR) – converting audio to text.

11.2 Core Principles

- Finance produces massive unstructured text from 10-Ks, 10-Qs, earnings calls, news, social media, etc.
- NLP compresses and structures text for trading, risk management, and compliance.
- Domain adaptation is crucial: generic sentiment tools often fail on finance jargon.
- Backtesting is necessary to avoid look-ahead bias in text-based signals.

11.3 Technical Details

- Typical NLP pipeline: ingest, pre-process, feature engineering, modeling, evaluation, and deployment.
- Sentiment approaches: lexicon-based (LM lexicon), machine learning models, and transformer-based models like FinBERT.
- SEC filings analysis: focus on MD&A, Risk Factors, Market Risk, Footnotes; compute tone, uncertainty, readability.
- Earnings call analysis: compare prepared remarks vs Q&A sections; analyze tone and evasiveness.
- News and social media: real-time ingestion, entity resolution, sentiment, novelty, and velocity metrics.

11.4 Important Distinctions

- Lexicon vs ML vs Transformer methods: rule-based vs data-driven vs deep contextual representations.
- Document-level vs sentence-level sentiment.
- NLP for trading vs NLP for compliance: speed vs auditability.
- Text sentiment vs numeric fundamentals: complementary signals.

11.5 Key Takeaways – Topic 11

- NLP scales analysis from pages of text into usable signals.
- You should be able to describe a basic NLP pipeline and where errors can arise.
- Finance-specific tools like LM lexicon, FinBERT, and F1 scores are important.
- Always connect NLP outputs to concrete decisions (trade, hedge, engage management).

11.6 Common Pitfalls – Topic 11

- Using generic consumer sentiment tools on finance text.
- Ignoring sarcasm, negation, and context.
- Not backtesting sentiment signals before using them in trading or risk systems.

Topic 12 – Credit Risk

12.1 Key Concepts & Definitions

- Credit Risk – risk of loss from borrower or issuer default.
- Probability of Default (PD) – chance of default over a horizon.
- Loss Given Default (LGD) – fraction of exposure lost when default occurs.
- Exposure at Default (EAD) – amount outstanding if default happens.
- Expected Loss (EL) – EL = PD × LGD × EAD.
- Revolving Credit – credit lines or cards with flexible balances.
- Installment Loans – fixed schedule loans such as auto or student loans.
- Debt-to-Income (DTI) – total debt payments divided by income.
- Loan-to-Value (LTV) – loan amount divided by collateral value.
- FICO Score (300–850) – standardized consumer credit score.
- Credit Bureaus – Equifax, Experian, TransUnion.
- Scorecard – points-based mapping from features to a risk score.

12.2 Core Principles

- Credit risk combines likelihood of default (PD) and loss severity (LGD).
- Traditional credit scoring uses logistic regression plus scorecards.
- ML models can improve predictive power but may reduce interpretability.
- Fair lending and privacy laws strongly shape data and model choices.

12.3 Technical Details

- Expected Loss formula: EL = PD × LGD × EAD.
- Traditional credit scoring inputs: bureau data, tradelines, utilization, inquiries, income, DTI, LTV.
- Application vs Behavioral scorecards: new borrowers vs existing accounts.
- ML models for credit: decision trees, random forests, SVMs, neural networks, ensembles.
- Explainability tools: SHAP values and LIME for local and global explanations.
- Key regulations: ECOA/Reg B, FHA, GDPR, Basel III, SR 11-7.

12.4 Key Takeaways – Topic 12

- Memorize EL = PD × LGD × EAD.
- Understand FICO components and scorecard logic at a high level.

- Know the trade-offs between logistic regression and ML ensembles.
- Recognize why interpretability and governance matter in credit risk.

12.5 Common Pitfalls – Topic 12

- Treating PD, LGD, and EAD as fixed numbers rather than state-dependent.
- Using protected attributes directly in models.
- Ignoring model risk management requirements.
- Assuming the most complex model is automatically best.

Topic 13 – AI in Capital Budgeting Decisions

13.1 Key Concepts & Definitions

- Capital Budgeting – deciding which long-term investment projects to accept.
- Net Present Value (NPV) – present value of inflows minus present value of outflows.
- Internal Rate of Return (IRR) – discount rate where NPV equals zero.
- Payback Period / Discounted Payback – time to recover initial investment.
- Profitability Index (PI) – present value of inflows divided by initial outlay.
- Weighted Average Cost of Capital (WACC) – weighted average of cost of equity and debt.
- Real Options – option-like flexibility embedded in projects.
- MAPE / RMSE – forecast accuracy metrics.

13.2 Core Principles

- Traditional capital budgeting uses discounted cash flow and NPV rules.
- AI enhances cash flow forecasting, risk analysis, scenario planning, optimization, and monitoring.
- Always align discount rates and risk metrics with project risk, not just firm averages.

13.3 AI-Enhanced Capital Budgeting

- Predictive cash flow modeling using ML models and rich feature sets (POS, CRM, ERP, macro data).
- Advanced risk assessment with Monte Carlo simulation, VaR, and CVaR of NPV.
- Use of GARCH or LSTM volatility models to link market volatility to cash flow and discount rate risk.
- Automated data pipelines: NLP to summarize industry news, APIs to pull macro and competitor data.

- Enhanced scenario planning: generative models propose plausible future scenarios and stress tests.
- Optimization and decision support: project portfolio optimization, explainable recommendations.
- Real options and post-investment monitoring using ML for optimal timing and anomaly detection.

13.4 Key Takeaways – Topic 13

- Be fluent with NPV, IRR, Payback, PI, and cash flow estimation.
- Understand how AI improves forecasting, risk modeling, and project selection.
- Know definitions of MAPE, RMSE, VaR, CVaR, and real options.

13.5 Common Pitfalls – Topic 13

- Comparing projects solely on IRR when NPV differs.
- Ignoring real options in staged projects.
- Overfitting ML forecasting models and not monitoring drift.

Topic 14 – AI and Financial Engineering (Derivatives)

14.1 Key Concepts & Definitions

- Derivative – security whose payoff depends on an underlying asset or index.
- Forward – OTC agreement to transact in the future at a fixed price.
- Future – exchange-traded forward with margin and daily mark-to-market.
- Option (Call/Put) – right, not obligation, to buy/sell at a strike price.
- Swap – exchange of cash flow streams (e.g., fixed vs floating).
- Exotic Options – barrier, digital, Asian, and other path-dependent options.
- No-Arbitrage / Law of One Price – identical cash flows imply identical prices.
- Risk-Neutral Valuation – pricing as discounted expected payoff under a risk-neutral measure.
- Black-Scholes-Merton (BSM) – classic option pricing model for European options.
- Greeks – sensitivities of option price: delta, gamma, vega, theta, rho.

14.2 Core Principles

- Derivatives are used for hedging, speculation, and arbitrage.
- Pricing is anchored by no-arbitrage and replication.
- AI acts as a fast, flexible approximation layer around traditional pricing frameworks.

14.3 AI & ML Enhancements

- Supervised learning models approximate complex pricing functions using training data from traditional models.
- Gradient boosting and other ML methods calibrate implied volatility surfaces.
- Deep neural networks act as neural surrogates for Monte Carlo pricing, offering large speedups.
- ML models improve volatility and risk forecasting by incorporating richer features and non-linearities.
- Sequence models like LSTMs and Transformers capture dynamics in return and volatility time series.
- NLP-derived sentiment and event detection feed into volatility and jump risk forecasts.

14.4 Key Takeaways – Topic 14

- Understand basic derivative types and their uses.
- Know high-level no-arbitrage and risk-neutral valuation logic.
- Explain how AI provides faster pricing, better volatility forecasts, and new trading signals.
- Recognize model risk and regulatory constraints in AI-driven derivatives trading.

14.5 Common Pitfalls – Topic 14

- Treating AI-based prices as correct without checking for no-arbitrage violations.
- Ignoring fat tails and jumps in return distributions.
- Using time-series cross-validation incorrectly and allowing look-ahead bias.

Quick-Reference Glossary (Selected)

- **Arithmetic Mean Return** – simple average of periodic returns.
- **Asian Option** – payoff based on average underlying price over time.
- **Beta (β)** – sensitivity of asset returns to market returns.
- **Black Box Model** – internal mechanics opaque; only inputs/outputs observed.
- **Black-Scholes-Merton (BSM)** – closed-form pricing model for European options.
- **Capital Budgeting** – evaluating long-term investment projects.
- **Capital Rationing** – choosing projects when capital is limited.
- **CAGR / Geometric Mean** – compounded average growth rate.
- **Correlation (ρ)** – normalized measure of linear co-movement (−1 to +1).
- **Credit Scorecard** – points-based credit scoring system.
- **CVaR (Expected Shortfall)** – expected loss given that VaR is exceeded.
- **DTI (Debt-to-Income)** – ratio of debt payments to income.
- **EAD** – Exposure at Default.
- **EL** – Expected Loss = PD × LGD × EAD.

- **F1 Score** – harmonic mean of precision and recall.
- **FICO Score** – standardized consumer credit score (300–850).
- **FinBERT** – transformer model fine-tuned for finance text.
- **Fog Index** – readability metric (years of education needed).
- **GARCH** – volatility model capturing clustering.
- **Greeks** – sensitivities of derivative price to drivers.
- **IRR** – discount rate where NPV = 0.
- **LGD** – Loss Given Default.
- **LIME / SHAP** – model-agnostic explainability tools.
- **MAPE** – Mean Absolute Percentage Error (forecast accuracy).
- **Max Drawdown (MDD)** – largest peak-to-trough loss.
- **Monte Carlo Simulation** – repeated random sampling to approximate distributions.
- **NLP** – Natural Language Processing.
- **NPV** – Net Present Value.
- **PD** – Probability of Default.
- **POS / CRM / ERP** – operational data systems feeding ML models.
- **Profitability Index (PI)** – PV inflows / initial investment.
- **Risk-Neutral Valuation** – pricing via expected payoff under risk-free drift.
- **Scorecard** – mapping from features to score in credit models.
- **Sharpe Ratio** – excess return per unit of σ.
- **Sortino Ratio** – excess return per unit of downside σ.
- **TF-IDF** – text representation emphasizing rare but important terms.
- **VaR** – Value at Risk (quantile of loss distribution).
- **WACC** – Weighted Average Cost of Capital.