Machine Learning for Retirement Planning

GORDON IRLAM

GORDON IRLAM is an independent retirement researcher in Los Altos, CA.
gordoni@gordoni.com

KEY FINDINGS

- The field of machine learning is evolving rapidly. Within machine learning, the sub-field of reinforcement learning appears to have the most relevance to retirement planning.
- Reinforcement learning is capable of handling the real-world complexity of financial planning, including the effects of income taxes, mean-reverting asset classes, time-varying bond yield curves, and uncertain life expectancies.
- For simple scenarios, machine learning delivers to within a few percent of the optimal solution. For more complex scenarios, whose optimal solution is unknown, machine learning is found to outperform other common approaches.

ABSTRACT: *Machine learning provides a new approach to solving problems in many fields. This article explores the use of machine learning to solve the retirement portfolio problem: deciding how much wealth to consume and how to allocate the remainder. After first reviewing existing approaches to the portfolio problem, this article looks in detail at the use of reinforcement learning. For simple financial scenarios where the optimal solution is known, reinforcement learning is found to deliver to within a few percent of the optimal solution. For more complicated financial scenarios, with no known optimal solution, reinforcement learning outperforms other common approaches. Reinforcement learning proves capable of optimizing highly complex financial models, including the effects of income taxes, mean-reverting asset classes, and time-varying bond yield curves, all of which other approaches cannot handle. Reinforcement learning appears to be the first fundamentally new approach to the portfolio problem in over 50 years.*

TOPICS: ***Retirement, big data/machine learning****

*All articles are now categorized by topics and subtopics. **View at PM-Research.com.**

The portfolio problem considered in this article involves deciding how much wealth to consume now and how to invest any remaining assets so as to maximize lifetime well-being. Despite its seemingly simple nature, real-world complexities make the solution challenging. Real-world complexities include factors such as labor income, Social Security, and other guaranteed income; income taxes; taxable, tax-deferred, and tax- free savings; mean-reverting asset classes, stock return volatility predictability; fat-tailed stock returns; time-varying bond yield curves; inflation; uncertainty in the true asset class mean returns given the finite historical record; leverage and borrowing constraints; uncertain life expectancies; and the availability of single premium immediate annuities (SPIAs) as a form of longevity insurance.

Several approaches touch on the portfolio problem: modern portfolio theory (MPT), Bayesian approaches, Merton's portfolio problem, stochastic dynamic programming, Monte Carlo trial and error, and rules of thumb. But all of these approaches are far from perfect. They either don't really solve the problem, are unable to handle the full complexity of the problem, are computationally too demanding, are time consuming to perform, or deliver far from the optimal solution.

Machine learning is a relatively new field that has had considerable success at addressing hard problems across many domains. To date, however, little work has involved exploring the use of machine learning for the portfolio problem.

After reviewing the traditional approaches to the portfolio problem, this article explores how machine learning can be applied to solve the portfolio problem. This is followed by two detailed examples. First, the article investigates a very simple scenario using machine learning compared with the exact results computed using stochastic dynamic programming, to good effect. Second, it looks at the results for a highly detailed financial model compared with rules of thumb, also to good effect. The article concludes with some caveats regarding the use of machine learning for retirement planning.

TRADITIONAL APPROACHES TO THE PORTFOLIO PROBLEM

Modern Portfolio Theory

Despite its importance to asset pricing, MPT is a non-solution to the portfolio problem. First, it says nothing about how much to consume. Second, its results depend on how much risk you should currently take. Unfortunately, the amount of risk you should currently take varies depending on your assets, guaranteed income, and possibly even age. And MPT is silent on the level of risk to take at each point.

Bayesian Approaches

Bayesian approaches incorporate into their analysis prior beliefs and uncertainties related to the system being modeled. To date Bayesian approaches have largely existed as refinements to MPT, and have some of the same issues.

Merton's Portfolio Problem

Merton's portfolio problem (Merton 1969) provides a precise mathematical solution to a highly simplified version of the portfolio problem. In Merton's portfolio problem there are a set of log-normally distributed asset classes, a risk-free asset, no leverage or borrowing constraints, no guaranteed income, and a fixed finite or infinite lifespan.

The lack of guaranteed income is particularly problematic. For many retirees, guaranteed income in the form of Social Security payments represents their largest asset. It is thus important that the effects of guaranteed income be included in any retirement plan.

Stochastic Dynamic Programming

Stochastic dynamic programming (Bellman 2003) is a computational technique that is capable of delivering the optimal solution to a version of Merton's portfolio problem that includes guaranteed income.

Unfortunately, attempting to extend stochastic dynamic programming to incorporate more than one or two additional real-world factors is computationally intractable.

Monte Carlo Trial and Error

The trial and error approach to the portfolio problem involves using a Monte Carlo simulator to explore different strategies and finding one that performs best.

The trial and error approach is both time consuming and samples only a very small subset of possible strategies. Typically, only strategies that are constant or simple linear functions of age, life expectancy, or portfolio size are considered. The trial and error approach also often fails to consider strategies that are responsive to other factors, such as guaranteed income and market conditions.

Rules of Thumb

Rules of thumb are often used when faced with the portfolio problem in real life. Examples of rules of thumb include following a glide path while working, a fixed asset allocation in retirement, an actuarial withdrawal scheme, and annuitizing by age 80.

Rules of thumb are a weak approach to the portfolio problem. They are highly unlikely to deliver close to the optimal result most of the time.

REINFORCEMENT LEARNING

All of the traditional approaches to the portfolio problem have severe limitations. It is time to consider a new approach.

The use of machine learning to perform asset allocation by itself was explored some time ago (Neuneier 1996, 1998). The addition of consumption occurred only slightly more recently (Sato and Kobayashi 2000). This early work made use of the machine learning concepts available at that time, such as Q-learning, rather than more modern algorithms, such as proximal policy optimization, that are available today.

Over the past 10 years, significant advances in machine learning have resulted in it being able to solve problems across a wide number of fields. More recently a branch of machine learning called reinforcement learning has been recognized as applicable to a broad range of problems in finance (Kolm and Ritter 2019). It is thus now reasonable to ask how reinforcement learning might fit in with retirement planning.

It was not a goal of this paper to see if reinforcement learning can learn retirement planning starting with a blank slate; that has been explored by prior research. Instead a major goal of this paper was to evaluate the performance of a carefully tuned reinforcement learning model. Thus the reinforcement model starts out with some sensible defaults, which the reinforcement learning algorithm is then allowed to fine tune. These defaults include an initial consumption estimate of wealth divided by retirement life expectancy, and an initial stock allocation that asymptotically approaches a fixed percentage as guaranteed income approaches zero. Complete details of the reinforcement learning model are provided elsewhere (Irlam 2020).

Reinforcement learning seeks to maximize the rewards obtained by an agent taking a sequence of actions in an observable environment. To cast the portfolio problem in reinforcement learning terms, imagine some entity (the agent) performing a sequence of asset allocation and consumption decisions against some financial model (the environment). At each time step, the agent gets to observe the current state of the financial model: age, portfolio size, guaranteed income, interest rates, and so on. Then based on this observation, the agent decides how much to consume and how to allocate assets, including possibly even purchasing SPIAs. The decided-upon action is taken as input to the environment and the financial model is advanced by one time step. And then the process repeats, with the agent performing another observation and deciding on the next action. The goal of the agent is to maximize aggregate well-being over time or, more technically, the sum of the utilities of consumption for each time period weighted by the probability of being alive.

Reinforcement learning involves a training phase, in which a neural network is trained on the financial model. This is followed by an inference phase in which the trained neural network can be used to predict a near optimal action to take for a given financial situation. The training phase typically takes a long time and involves a large amount of computer power, while inference is very fast and requires relatively modest computer resources.

Reinforcement learning doesn't deliver the optimal solution. The results might be very good, but they are not optimal. Training the neural network for a longer period of time with a smaller step size, using a larger neural network, or increasing the fidelity of the financial model observations can reduce the distance between the optimal solution and the reinforcement learning solution.

Improvements in machine learning have gone hand in hand with the use of larger and larger data sets. The use of reinforcement learning for the portfolio problem is no exception, requiring tens to hundreds of millions of observations to train. Fortunately, it is possible to use a financial simulator to generate this data on the fly. At first, the neural network has little idea of the correct actions to perform or the expected rewards, but after millions of time steps, it builds up a model of the correct action to take and the expected rewards from doing so.

Training a neural network to handle a single financial scenario is relatively easy and delivers very good financial results. There is a problem with this approach though. Training a neural network for each financial client would be both costly and time consuming. Training a generic neural network capable of inferring a range of scenarios is a harder problem and is the focus of this article. Scenarios might differ in terms of the initial amount of assets present, guaranteed income, client health, availability of a 401(k) scheme, and so on.

VALIDATION OF THE REINFORCEMENT LEARNING APPROACH

Before applying reinforcement learning to the fully general retirement portfolio problem, it is worthwhile to compare the performance of reinforcement learning with some other techniques known to deliver the optimal strategy. For this analysis, stochastic dynamic programming was used. On account of the limitations of stochastic dynamic programming, the financial model used was highly simplified.

An uncertain mortality was used: a female with a starting age of 67 and a probability of death specified by the Social Security Administration's AS 120 cohort life table. To reflect the fact that financial planning clients are typically in better than average health, a constant amount was subtracted from the age so that the initial life expectancy was three years longer than it otherwise would have been.

The financial model provided for $20,000 of Social Security income per year. For this comparison, taxes were ignored.

Several initial portfolio sizes were considered. Real stock and bond returns were assumed to be log-normally distributed and non-correlated. Stock returns were taken from the global historical record specified in the 2020 Credit Suisse yearbook (Dimson et al. 2020). The bond return was the 15-year Treasury real yield curve on December 31, 2019, of 0.3%, while the bond volatility was taken from the Credit Suisse yearbook. For the validation model, no attempt was made to account for the standard errors in measurement of the returns data.

Utility maps an amount of consumption to the amount of satisfaction, or well-being, derived from that consumption. The relationship between consumption and utility is non-linear because satisfaction saturates at higher levels of consumption. Utility of consumption was determined using a mathematical constant relative risk aversion utility function. Relative risk aversion (RRA) coefficients of 1.5, 3, and 6 were used.

A single generic neural network was trained for each RRA coefficient. Each generic neural network was trained using a broad range of portfolio sizes and guaranteed incomes.

The performance of a strategy for some scenario is measured by computing its certainty equivalent (CE). The CE is the constant fixed consumption with the same utility as the average variable utility experienced under the strategy.

EXHIBIT 1

Reinforcement Learning Compared to Optimal Results for a Simple Financial Model

RRA	Initial Wealth	Reinforcement Learning CE	Optimal CE	Relative Performance
1.5	$200,000	$32,828	$32,881	99.8%
1.5	$500,000	$50,157	$50,221	99.9%
1.5	$1,000,000	$77,933	$78,045	99.9%
1.5	$2,000,000	$132,100	$132,451	99.7%
3	$200,000	$31,598	$31,636	99.9%
3	$500,000	$46,262	$46,324	99.9%
3	$1,000,000	$68,651	$68,828	99.7%
3	$2,000,000	$110,591	$111,265	99.4%
6	$200,000	$29,985	$30,091	99.6%
6	$500,000	$41,105	$41,392	99.3%
6	$1,000,000	$57,009	$57,801	98.6%
6	$2,000,000	$85,936	$87,696	98.0%

Note: Results are for a retiree, age 67, in good health, and Social Security of $20,000 a year.

The performance of the trained models is shown in Exhibit 1. For all of the scenarios considered, reinforcement learning delivered 98% or more of the optimal value.

Reinforcement learning knows nothing of log-normally distributed returns and the cohort life expectancy tables. The very good performance of reinforcement learning on the simple financial model suggests it might be able to also deliver good results on a richer financial model.

A RICH FINANCIAL MODEL

A rich financial model was developed as specified in the following. It isn't necessary to understand all the precise details of the financial model. The important point is the model comes far closer to representing the real world than the simplified financial models of traditional approaches.

Mortality

Mortality is as specified by the female AS 120 cohort life table with a variable health-related adjustment applied to the initial starting age.

Stock Model

The starting point for stock market returns is the reported global mean and standard deviation of returns from the 2020 Credit Suisse yearbook. The reported standard error, or uncertainty in the true mean, is applied to give a different simulated mean return for each stochastic sample sequence.

Shiller makes the case that stocks at times trade above or below their rational valuations (Shiller 2015). This is implemented by reducing or increasing the expected mean return by 1% for every 10% stocks are overvalued or undervalued. A precise number for how much of a correction to apply is difficult to estimate, but this adjustment is of the right magnitude. It is felt that applying this correction gives a more accurate model of the stock market than assuming just a simple random walk. The fair price against which stocks are seen to be over or undervalued does follow a random walk.

A monthly GJR-GARCH volatility model is also applied so that volatility has a degree of predictability. Historical returns from the S&P 500 Index are used to create the stochastic shocks of the volatility model. As a result, the stock returns display skew and kurtosis, or fat-tails.

Both the relative price level and the current volatility of stocks are able to be observed by the agent. This makes it possible for the agent to engage in tactical asset allocation. Reflecting the difficulty of knowing whether stocks are overvalued in the real world, a 15% standard deviation noise factor is applied to the observed relative stock price level.

Bond Model

The bond model is a Hull–White bond model. This means there is a yield curve and the shape of the yield curve varies over time non-linearly in response to changes in the short interest rate. The short interest rate is described by a random process with a standard error value given from the Credit Suisse yearbook. The neutral yield curve is intended to be representative of the present era and is based on the average Treasury yield curve from 2005 to 2019.

Both real and nominal yield curves are generated, with the difference between the two taken to represent the inflation rate. A corporate bond yield curve is also generated by applying the average of the Bank of America Merrill Lynch option-adjusted spreads for 1997 to 2019 to the nominal yield curve. The corporate bond yield curve is currently only used as an input to a SPIA pricing model. SPIAs are not used for the results presented in this article.

The duration of bonds is selected by the agent as a sub-component of the asset allocation decision. In practice, the agent was found to show a strong preference for long duration bonds. Real bonds were found to outperform nominal bonds, and so based on this, the ability to allocate assets to nominal bonds was disabled.

Returns are intended to be indicative of a bond fund. At the end of a time period, the return is computed from the interest rate plus the change in rate and the duration.

Tax Model

The tax model is based on the 2020 US tax code. Portfolio wealth and income are both divided into tax-free, tax-deferred, and taxable buckets.

In addition to regular income taxes, there are capital gains taxes, the net investment income tax, and state, local, and property taxes. Social Security is taxed specially. Capital gains are computed using the average cost tax basis method. State, local, and property taxes are crudely represented as equal to 11% of income above the standard deduction. This is based on an average of the state, local, and property tax burdens. Improving much beyond this would require exploring the details of the tax regimes of every state. This would be a major undertaking.

IRA and possibly 401(k) contributions are allowed with catch-up provisions for older workers. The agent is able to observe whether a 401(k) is available.

Guaranteed Income Model

Multiple sources of guaranteed income are possible. Each source of guaranteed income may be real or nominal, and tax free, tax deferred, or taxable. The income may be present over a fixed number of years or exist until death, and it may grow by a fixed amount each year.

PERFORMANCE IN RETIREMENT

Retirement planning is easiest for a low coefficient of risk aversion and a high ratio of guaranteed income

EXHIBIT 2

Performance of Various Strategies for a Detailed Financial Model

Withdrawal Scheme	Asset Allocation	CE
Percent Rule	Fixed	$69,058
Guyton's Rule 2	Fixed	$69,902
PMT, Dynamic Life Expectancy	Fixed	$72,805
Target Percentage Adjustment	Fixed	$73,448
Guyton–Klinger	Fixed	$80,884
Extended RMD	Fixed	$83,686
PMT, Fixed Life Expectancy	Fixed	$85,810
Blinded Reinforcement Learning	Dynamic	$88,802
Reinforcement Learning	Dynamic	$91,215

Notes: Results are for a retiree, age 67, in good health, RRA 6, Social Security of $20,000 a year, and a $2,000,000 investment portfolio. PMT is the constant amount required to deplete the portfolio over the time period. Extended RMD is the proposed IRS required minimum distribution for 2021 extended back to age 67.

to portfolio wealth. Conversely, it is most challenging for a high coefficient of risk aversion and a low ratio of guaranteed income to portfolio wealth. Therefore, to showcase the performance of reinforcement learning, an RRA of 6, $20,000 of taxable Social Security, and a $2,000,000 portfolio was used. Once again, a 67-year-old female retiree in good health was considered.

Exhibit 2 shows the results for the retiree in the rich financial model for a number of rules of thumb and reinforcement learning. Most of the rules of thumb have such parameters as asset allocation, assumed life expectancy, and assumed rate of return. Several days were spent performing a trial and error search to determine the optimal parameter values for all of the rules of thumb. The reader is referred to the literature for details on many of these rules (Bengen 1994; Guyton 2004; Guyton and Klinger 2006; Zolt 2013). PMT takes its name from an Excel function and is the constant amount required to deplete the portfolio over the indicated time period assuming it is growing at some constant rate. Extended RMD is the proposed IRS required minimum distribution table for 2021 extended back to age 67. Blinded reinforcement learning is reinforcement learning blinded to observation of the relative stock price, volatility, and short interest rate. All of the rules of thumb were extended in the natural fashion to handle the presence of guaranteed income and taxes. Reflecting the long duration typically selected by reinforcement learning, a 20-year bond duration was used for the rules of thumb.

The rankings of the different rules of thumb should not be seen as an overall comparison between them. That would require comparing how they perform across many different scenarios. Indeed, for a $1,000,000 investment portfolio quite different rankings are obtained, but reinforcement learning still outperformed all the rules of thumb. Note that some of the optimal parameter values are far from what are commonly used. For example, for this scenario, the optimal fixed life expectancy for PMT was 50 years, which is far higher than the 30- or 35-year retirement life expectancy commonly used by financial planners. Using a fixed life expectancy of 35 years, PMT delivers a CE of $48,527; lower than any other rule. The takeaway should be reinforcement learning outperforms common rules of thumb and does so without requiring a tedious trial and error search of different rules and parameter values.

The small difference between blinded reinforcement learning and reinforcement learning that is able to observe the stock and bond prices points to the challenges of successfully performing tactical asset allocation.

DISCUSSION

Many problems in financial planning have either a definite answer or there exists a framework for figuring the answer out. For a long time, this was only weakly true of the portfolio problem. Stochastic dynamic programming was unable to handle the real-world complexity of the portfolio problem. On the key issue of asset allocation, people often fall back on doing what they feel the most comfortable with in the short term, not what will produce the best outcome for them in the long term. This need no longer be the case. Reinforcement learning promises to provide scientifically valid answers to a field long dominated by alchemy.

One of the potential problems with reinforcement learning is the black box nature of the recommendations. Reinforcement learning provides no insight into why a particular asset allocation is recommended, other than the fact that it delivered good results over millions of simulated time periods.

A deployment difficulty for reinforcement learning stems from the fact that a 10% change from the optimal asset allocation has a very small effect on the CE; less than 0.5%. As a result, different trainings of a neural network for the same financial model might produce different recommendations. This doesn't really matter in the grand scheme of things, but it needs to be explained to the client why the same financial parameters produce different recommendations based solely on which neural network model was used.

Reinforcement learning may be expected to perform poorly for scenarios that are outliers and only infrequently encountered by the training process. Caution should probably be used in interpreting the results in these cases or from specific models trained to handle just these individual scenarios.

There is a risk of people seeing reinforcement learning for financial planning as more than it is simply because it is capable of being wrapped up in the label "artificial intelligence." Reinforcement learning appears to do an excellent job of solving the portfolio problem, but it is not all knowing, and it is only as accurate as the financial model on which it is built.

Reinforcement learning doesn't solve the difficult personal decision of what is an appropriate coefficient of relative risk aversion. Thus, it may appear things are no better than the case with MPT where we are stuck wondering how much risk to take and thus what is the appropriate asset allocation to use. This is not the case. First, the use of a consumption risk aversion framework recasts the problem from one of how do we feel about changes in investment portfolio value to the more meaningful how do we feel about changes in retirement consumption. Second, the use of a consumption risk aversion framework makes the results self-consistent over the lifecycle. There is a single subjective factor to determine, instead of having to determine subjective factors at every point of the lifecycle.

AVAILABILITY

The model described here is available at www.aiplanner.com. It requires no setup and is easy to run. You answer a few high-level questions, and in return it produces financial recommendations and performs a Monte Carlo simulation of what to expect in the future. AIPlanner also includes features for pre-retirement planning and annuitization using SPIAs. AIPlanner is described in greater detail elsewhere (Irlam 2018, 2020).

CONCLUSION

Reinforcement learning is the first fundamentally new approach to the portfolio problem in more than 50 years. For simple scenarios with a known solution, it delivers to within a few percent of that solution. And for complex scenarios with no known solution, it is found to outperform commonly used rules of thumb. For a long time, retirement planners have been forced to guess at how to allocate assets and how much wealth to consume; with reinforcement learning, this need no longer be the case.

ACKNOWLEDGMENTS

The author is grateful to Phil DeMuth, William Selden, and Joe Tomlinson for feedback.

REFERENCES

Bellman, R. 2003. *Dynamic Programming*. Dover Publications.

Bengen, W. P. 1994. "Determining Withdrawal Rates Using Historical Data." *Journal of Financial Planning* 7 (4): 171–180.

Dimson, E., P. Marsh, and M. Staunton. 2020. *Credit Suisse Global Investment Returns Yearbook 2020*. Zurich: Credit Suisse.

Guyton, J. T. 2004. "Decision Rules and Portfolio Management for Retirees: Is the "Safe" Initial Withdrawal Rate Too Safe?" *Journal of Financial Planning* 17 (10): 54–62.

Guyton, J. T., and W. J. Klinger. 2006. "Decision Rules and Maximum Initial Withdrawal Rates." *Journal of Financial Planning* 19 (3): 48–58.

Irlam, G. 2018. "Financial Planning via Deep Reinforcement Learning AI." Working paper, SSRN: http://ssrn.com/abstract=320713.

——. 2020. "Multi Scenario Financial Planning via Deep Reinforcement Learning AI." Working paper, SSRN: http://ssrn.com/abstract=3516480.

Kolm, P. N., and G. Ritter. 2019. "Modern Perspective on Reinforcement Learning in Finance." Working paper, SSRN: http://ssrn.com/abstract=3449401.

Merton, R. C. 1969. "Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case." *The Review of Economics and Statistics* 247–257.

Neuneier, R. 1996. "Optimal Asset Allocation using Adaptive Dynamic Programming." In *Advances in Neural Information Processing Systems 8*, proceedings from Neural Information Processing Systems 1997, pp. 952–958. https://papers.nips.cc/paper/1121-optimal-asset-allocation-using-adaptive-dynamic-programming.

——. 1998. "Enhancing Q-Learning for Optimal Asset Allocation." In *Advances in Neural Information Processing Systems 10*, proceedings from Neural Information Processing Systems 1997, pp. 936–942. https://papers.nips.cc/paper/1427-enhancing-q-learning-for-optimal-asset-allocation.

Sato, M., and S. Kobayashi. 2000. "Variance-Penalized Reinforcement Learning for Risk-Averse Asset Allocation." In *International Conference on Intelligent Data Engineering and Automated Learning*, pp. 244–249. Springer.

Shiller, R. J. 2015. *Irrational Exuberance: Revised and Expanded Third Edition*. Princeton University Press.

Zolt, D. M. 2013. "Achieving a Higher Safe Withdrawal Rate with the Target Percentage Adjustment." *Journal of Financial Planning* 26 (1): 51–59.

ADDITIONAL READING

Retirement Income Research: *What Can We Learn from Economics?*
Gordon Irlam and Joseph Tomlinson
The Journal of Retirement
https://jor.pm-research.com/content/1/4/118

ABSTRACT: *Research on retirement income planning has, for many years, followed two separate tracks. Financial planning practitioners have developed guidelines for withdrawals from savings and asset allocation by testing and fine-tuning various rules of thumb. Economists have applied a different approach, based on life-cycle finance, aimed at maximizing the utility of lifetime consumption, using dynamic programming techniques to optimize retirement withdrawals and asset allocations. This article seeks to use a non-mathematical explanation to help financial planners and others who are not familiar with the economics approach to develop a conceptual understanding of how life-cycle finance and dynamic programming can be applied to retirement income planning. It also compares the performance of optimized recommendations produced by dynamic programming with various rule-of-thumb strategies that have been popular in the planning literature. It attempts to demonstrate the power that the economics approach can bring to improving retirement income planning and argues that more communication between economists and practitioners can help open up new approaches to research and improved practical applications.*