ISSN: 1803-3814 (Printed), 2571-3701 (Online)
https://doi.org/10.13164/mendel.2019.1.057

EVOLVING PREDICTIONS FOR EXECUTIVE PAY FEATURES IN BOARD NETWORKS

Ami Hauptman[✉]**, Amit Benbassat and Rosit Rosenboim**
Computer Science and Economics Departments, Sapir College, Israel
amih@sapir.ac.il[✉], amitbenb@mail.sapir.ac.il, rositl@mail.sapir.ac.il

Abstract
Numerous recent studies in finance literature have shown that board networks are an important inter-corporate setting, influencing corporate decisions made by the board of directors, for example the determination of executive pay features. In this paper, we evolve predictors for the existence and adoption of several important pay features among S&P1500 companies, over the period 2006–2012. We use data from five well-known financial databases, including hundreds of variables containing both director-level and firm-level data. We present two approaches for predicting executive pay features. The first approach is based on a Genetic Algorithm (GA) used to evolve predictors based on weighted vectors of the predicting variables, providing relatively easy to understand prediction rules. The second approach employs Genetic Programming (GP) with sets of functions and terminals we devised specifically for this domain, based on contemporary research in finance. Thus, the GP approach explores a wider problem space and allows for more complex feature combinations. Experiments using both methods attain high quality prediction results, when compared to previous results in finance research. Additionally, our model is capable of successfully predicting combinations of pay features, compared to standard empirical models in finance, under various experimental conditions.

Keywords: *finance, genetic algorithm, genetic programming, prediction, pattern recognition.*

Received: 13 April 2019
Accepted: 30 May 2019
Published: 24 June 2019

1 Introduction

Ample research in the field of corporate finance explores decisions made by the board of directors and the factors influencing them [8, 9, 3, 10, 28]. Besides traditional characteristics, such as company accounting, financial data and personal directors data, recent studies have found that board networks also have significant impact upon corporate decisions [7, 5, 6, 11, 12]. Board networks are graphs describing the connections between firms. Each node in the network represents a company and each edge between two nodes represents at least one common director. One of the responsibilities of the board of directors is setting executive compensation. The influence of board networks over decision regarding executive *pay features* (*PF*s) has been well documented in recent literature [1, 21, 19, 21, 26].

To predict executive pay features, most corporate finance empirical researchers cited above apply *regression analysis* techniques to data extracted from several standard databases (for example *Center for Research in Security Prices* (CRSP) and *Compustat* for financial and accounting firm's data). While regression analysis has been widely successful in this field, this method has some important limitations: (1) Only one pay feature (the dependent variable) can be predicted by a given regression model. (2) Interpreting the results of such models may be difficult when the variables used for prediction are co-linear, which is often the case with financial data (e.g. variable describing firm size and firm profitability tend to be highly correlated). (3) Regression models are only applicable to testing given connections between variables, and not to finding new connections, and thus are not suitable for tasks such as selecting variables for prediction. (4) Such models typically examine only linear relations between variables, and cannot be used for example to test more complex relations.

Despite the importance of this problem, very little work has been done to apply machine learning, evolutionary algorithms, or related methods, to this domain (see Section 2). Moreover, we could not find any work in the literature regarding the application of evolutionary methods for predicting pay features.

In this work, we apply two evolutionary methods to this problem: Genetic Algorithms (*GA*) and Genetic Programming (*GP*). While GA-based predictors represent weights for predicting variables, GP-predictors are more complicated, adding Boolean conditions and several different operators to the mix. We evolve and test our predictors with two datasets which we constructed using variables extracted from five standard financial databases, including hundreds of variables containing both director-level and firm-level data. We also implemented a standard regression model to compare to our evolved predictors, using standard tools in finance.

Our evolved predictors not only perform well on various prediction tasks, yielding highly accurate predictions under most experimental conditions, but also present several important advantages over regression methods commonly used in this field including the ability to successfully predict multiple pay features with the same predictors and also being agnostic to co-linearity in the predicting variables.

The contributions of this paper are as follows:

1. This is the first reported successful attempt to apply evolution to predicting Pay Features, an important problem in finance research.

2. Our evolved predictors are not only highly accurate, but can also predict multiple Pay Feature simultaneously and cope with co-linearity, which standards regression method cannot.

3. The model described herein can easily be applied to further economics research–by evolving predictions for any desired set of variables.

The rest of this paper is organized as follows: Section 2 presents previous work relevant for this research. In Section 3 we describe the methods used in our experiments. We put our best results to date in Section 4. Finally, we draw conclusions and point the way for future research in Section 5.

2 Previous Work

As mentioned above, while there is ample research in the field of economics regarding board networks, and specifically applying them to predict PFs, there is very little research in which evolution, or learning, is applied to this problem. We hereby survey the relevant literature.

2.1 Predicting Pay Features in Finance Research

As mentioned above, the influence of board networks over decisions regarding executive pay features has been well documented in corporate finance literature. We hereby survey some important works. Bizjak et al.[5] examine the influence of board networks on the use of options backdating in executive pay. They find that a firm is more likely to backdate its option if it is connected to another backdating firm through board networks. Similarly, Davis [11] and Davis and Greve [12] examine the spread of poison pill and golden parachutes through board networks.

In a more recent study, Bouwman [6] finds that not only are firms likely to select directors from firms with similar pay practices but these directors also influence the firm's governance to move toward the practices of their firms. Hallock [19], Barnea and Guedj [1]) and Renneboog and Zhao [26] all find that firms with more connected directors award their CEOs higher compensation.

2.2 Machine Learning Methods Applied to This Problem

Several works in the field of machine learning and data mining have been applied to this problem. For example, Wong et al.[29] examine analyze a data sample of 725 large U.S-based public companies, and using Exponential Random Graph Modeling Techniques demonstrate that common directors between firms (dubbed *director interlocks*) are positively linked with similarities in executive pay features.

In related work, Gygax et al.[18] apply stochastic network techniques to investigate the relation between common directors, common compensation consultants and industry membership, and conclude that package similarity comes through sharing directors and industry membership.

2.3 Related work with Evolutionary Methods

The evolutionary approach is well used in economics research. For example, Biethahn and Nissen [4] used EAs in order to generate economic simulations and models. Harrald and Kamstra [20] used evolved Artificial Neural Networks as tools for combining stock market forecasts. For further details, see recent survey by Nissen [25].

It is important to note that this is the *first* reported attempt to apply EA methods to predicting pay features.

3 Method

We hereby discuss our method. After explaining how we constructed our databases, we move to describing the process of evolving our two types of predictors–with GA and GP. Finally, we depict our experimental setting for all different experiments, including our measures of success.

Table 1: 6 Basic data variables

Variable name	Return type
log(`Total assets`)	The natural logarithm of company's total assets
`Market to book`	The ratio of market value of total assets to book value of total assets
`Debt/Total assets`	The ratio of the sum of long-term and short-term debt to the book value of total assets
`Return On Assets`	The ratio of net income total assets
`SIC`	Standard classification for industries.
`DMC`	The number of compensation consultants used by the firm

3.1 Variables and Datasets

In order to evolve our predicting models, we processed data from 5 standard databases, commonly used in finance research: *Center for Research in Security Prices (CRSP)* [14] and *Compustat* [27] (for financial and accounting firm's data), BoardEx (`www.boardex.com`)–for director personal data and networks and both *ExecuComp* and *Incentive Lab*–for compensation data.

Using these databases, we obtained data regarding 1500 firms (included in the S&P1500 index) [1], during the period 2006-2012 and contains 10,168 company-year observations and 96,794 director-year observations. Company level data also contain 5 variables representing important unique *pay features* (PFs), namely: Option, Stock, Accounting, Absolute and Relative. These variables identify firms that award stock options to their CEO, that link executive pay to a stock price metric, accounting metric, and absolute or relative performance metric respectively. Values of these variables of either 1 (representing that the PF was given at that year) or 0 (the PF was not given).

As our goal in this work is to evolve models for predicting PFs, we selected and analyzed approximately 300 variables from the aforementioned databases, containing information relevant to predicting PFs. Variables we selected correspond to groups of variables described in major works in this field (see, for example: [3, 5, 26]).

Using the analysis described below, we constructed two datasets for evolution: one with 33 values per row (dubbed *BASIC*) and the second database (dubbed *EXTENDED*) extending the first one with an additional 30 values per row. Each dataset we constructed contained 10,168 rows, corresponding to company-year observations mentioned above: the values in each row represent predicting and predicted variables for a given company, at a given year.

The **BASIC** database (33 values per row) contained the following for each row:

1. Company identification and fiscal year.
2. 6 basic variables which are commonly used in corporate finance research (e.g. [5, 26]) and describe major aspects of the firm, for example firm size, firm profitability and firm growth (see Table 1).
3. 10 PFs–one value for each of the 5 PFs for the given year, and 5 values for the next year.
4. 15 graph variables conveying information regarding the board network. Each variable is PF-specific, and related to the number of neighboring companies (i.e. companies with a common director at a given year), that used the same PF in the previous year: 5 variables count the number of such neighboring companies (one for each PF), 5 variables count the number of common directors with neighboring companies and 5 variables compute a logarithmic function of the number of neighboring companies (for more details–see [5, 1, 26]).

The **EXTENDED** database (63 values per row) contains all the variables of the BASIC group, with an additional 30 variables, which are typically used in finance literature (for example [3, 5, 6, 26]). These variables contain more detailed information regarding the given company as well as the board of directors, for example–the average age within the board; information regarding the CEO–namely: gender, age and salary; graph centrality [17] of the company and more. For brevity purposes, we cannot provide a full descriptive list of our variables. However, as the purpose of constructing the second database was mainly to test how our method scales with the number of variables, the exact details of variables used here are of lesser importance. [2]

We use the two datasets described above in various experimental settings, as explained below. It is important to note that not all information in a given data row is used as an input to our predicting models for every prediction. When our models are used to predict only one PF, only variables pertaining to that specific PF are given as input to the model. This matter is further described in Subsection 3.6.

[1] This is a stock market index of US stocks made by Standard & Poors which includes all stocks in the S&P500, S&P400 MidCap, and S&P600 SmallCap. This index covers 90% of the market capitalization of U.S. stocks.

[2] In future research we plan to study the effects of more specific groups of variables.

Table 2: Function nodes. Parameters types are: F_i: Floating-point, B_i: Boolean, V_i: Var, TR_i: Test Result. When possible, $Multiple$ input parameters of the *same* type are abbreviated with a digit (e.g. $3F$, $2B$, etc.).

Node name	Type	Return value
AND(2B) , OR(2B), NOT(1B)	B	Logical AND , OR , NOT
MIN3(3F), MAX3(3F), MIN4(4F), MAX4(4F)	V	Minimum/Maximum value among parameters
IF_VAR_BG_X(V_1,F_1)	B	If $V_1 > F_1$ then True else False
IF_VAR_BG_VAR(V_1,V_2)	B	If $V_1 > V_2$ then True else False
IF_A_BIGGER_B_0(3F)	F	If $F_1 > F_2$ then F_3 else 0.0
IF_A_BIGGER_B_1(3F)	F	If $F_1 > F_2$ then F_3 else 1.0
IF_A_BIGGER_B_0_1(F_1,F_2)	TR	If $F_1 > F_2$ then 0.0 else 1.0
VARS_EQ_CLOSE(V_1,V_2)	B	If $\|V_1 - V_2\| < 0.1$ then True else False
VARS_FAR(V_1,V_2)	B	If $\|V_1 - V_2\| > 0.5$ then True else False
IF_THEN_ELSE(B_1,F_2,F_3)	F	If B_1 then F_1 else F_2
IFTE_BIGGER_AB(V_1,V_2,F_1,F_2)	F	If $V_1 > V_2$ then F_1 else F_2

After attempting several methods, we normalized our variables using percentile normalization [22]. Apart from avoiding dominance of variables with larger values when performing predictions, this technique has several advantages in our experimental setting, especially in the GP setting (see Section 3.3), when comparing variables to constants.

Both GA and GP were implemented with the DEAP Python framework (see [15]).

3.2 GA

We evolved the first type of predictors using a standard GA. Individuals represented weights for the predicting variables: both the importance of each variable (weight magnitude) and the directional effect on prediction (the sign of the weight). Thus, each prediction is obtained by receiving a value-row of a given dataset as input, and returning a numerical prediction for next year's PF.

When predicting a single PF value for a given data row, we extract all non-PF variable values from that row, along with PF-specific values, perform the dot product with the set of weights and attain a single predicted value. Since the goal is to predict whether the PF would be either 1 or 0 (discrete predictions), and predicted values are continuous, we set the threshold for a positive prediction at 0.5 and above[3]. As follows, predictions are linear combinations of the input variables[4]. A constant feature was also added.

For experiments in which we predict more than one PF, however, *the same* weight-vector can be used to predict *all* relevant PFs for a given row. To accomplish this task, we repeat the process described for predicting one PF-for all given PFs, and obtain several predicted values.

3.3 GP

Predictors of the second type were GP trees with Strongly Typed Genetic Programming ($STGP$). Four Node types we used were: *Float*—floating point values, *Boolean*, *Var*—feature variables and *Test Result*—representing comparison results with other values.

We implemented a few basic domain-independent terminal nodes, as well as a variable nodes for each feature variable. Terminal nodes included: *ERC*—Ephemeral Random Constants, *True*, *False* and *Var*—feature variables. For function nodes we used arithmetic and logic functions, as well as several control flow functions. Functions are detailed in Table 2.

As stated above, we used percentile normalization. This method permits direct comparisons between variables which originally contained different values, since we are actually comparing their percentiles. It is also possible to directly compare all variables to constants (and to ERCs) in the range $[0, 1]$, thus giving rise to more compact models.

3.4 Fitness Function

The fitness functions we used measure the success in obtaining correct values for prediction. Since PF values are always 1 or 0 (i.e. utilized or not), we used several loss functions to measure prediction correctness:

1. *RMSE*: Root Mean Squared Error - measuring the average distance of each prediction from the correct value for the example at hand.

[3]This is commonly done in statistical learning systems such as Logistic Regression [22], where continuous prediction values are interpreted as the *probabilities* for the positive class

[4]Still, some of the variables represent non-linear functions-see Section 3.1

2. *LogLoss*: This measure, sometimes called The Cross-entropy Loss function (see for example [24]), is often used in prediction tasks, as it allows for large (non-linear) penalties for wrong predictions.

3. *PercentWrong*: Percentage of wrong predictions (or $1 - Accuracy$). This is the standard measure for prediction tasks in data mining and related methods (see [22])

Note that although our ultimate goal is only to minimize the third measure, that measure is actually *discrete*, since it considers each prediction as either right or wrong. Thus, in order to guide evolution in the right direction, we also added the first two measures, which are continuous.

Since some measures (especially *PercentWrong*) are not suitable for highly skewed populations[5], which occur in some of our experiments, we selected an equal number of cases from each predicted value, with all evaluations and tests.

Using all three measures, we get the following formula to minimize:

$$Fitness = K_1 * RMSE + K_2 * PercentWrong + K_3 * LogLoss$$

Empirically, we set $K_2 > K_1$ and $K_2 > K_3$, since both *RMSE* are *LogLoss* are typically more prominent during the first stages of evolution, and then gradually decreases as individuals begin making more accurate predictions. However, *PercentWrong* decreases much more slowly, and is more important during the latter stages of each experiment. We finally settled on the values $10, 1000$ and 50 for K_1, K_2 and K_3, respectively.

We split the data to train set (80%) and test set (20%)–randomly for each experiment. There are approximately 2000 examples for each PF, out of which we randomly draw 200 samples for each fitness evaluation, and calculate the 3-factor fitness described above.

3.5 GA and GP Parameters and Environment

As stated above, we use the *DEAP* environment for all experiments, with the following parameter configurations, obtained empirically: Population size 250–500, generation count 150–400, and genetic operator probabilities: reproduction 0.5, crossover 0.4, mutation 0.1. For the GP experiments we used tree growth mutation from tree-node chosen uniformly at random, and One-way crossover from tree-node chosen uniformly at random.

3.6 Experimental Settings

We now describe various characteristics for each of the experiments we conducted for evolving predictors. Each single experimental setting is devised by selecting a *single* value out of *each* of the following:

GP or GA the Evolutionary Algorithm used to evolve predictors.

Basic or Extended the dataset to used for evolution. It determines the size of the individuals: for the Basic dataset GA-individuals are evolved with a smaller number of weights (relative to the Extended dataset) corresponding to the smaller number of values in each row, and GP individuals contain less *VAR* terminals.

Predict or Adopt while in some experiments we predict PFs in the general case, following the work of [5], in other experiments we predict the *adoption* of PFs. Under this condition, we only use data rows in which the current PF value is 0, and predict whether the firm will *adopt* that PF for the following year. While prediction under such conditions is still conducted in the same manner, there are substantially less relevant training examples, since PFs tend to remain constant approximately 70% of the time. We further address this issue in 4.

Single PF, two PFs or all PFs The number of PFs in each experiment is either 1, 2 or 5 (all PFs). We either attempt to accurately predict only a single PF, or evolve predictors for all PFs. Evolving two PFs is a special condition only for the Adopt experiments (see above), due to a limitation of the dataset we used–there are significantly less suitable examples for evolution, for all but two PFs: Option and Absolute. Hence, a group containing these two PFs is the only combination of PFs applicable to Adopt experiments (see Table 5). Note that the same structure for predictors is in all cases (see explanations in sections 3.2, and 2).

Only results for the Options PF are reported herein, as results in other experiments we conducted with single PFs were highly similar.

To quantify the success of our predictors we used two measures. The first was *RMSE*–as described in Section 3.4. The second was Accuracy, or $(1 - PercentWrong)$ – representing the percentage of correct predictions.

For each different setting we performed 20 experiments, on average. Section 4 reports the best results obtained for each experimental setting.

[5]This is one reason for not choosing more common measures such as Precision and Recall.

Table 3: Option Pay-Feature Table. Method: GA, GP or REG is the predictor type. Params indicates the database used (Basic - 33 values per row, Extended - 63 values per row). Predict: predicting generic Pay-Features values, and Adopt: predicting change from 0 to 1 for a given Pay Feature. ACC - Training accuracy (with Test accuracy in parentheses), except for Regression where there is no test. RMSE - Root mean squared error. Best are marked in bold. This is true for the current and for all subsequent tables.

Method	Params	Predict			Adopt	
		ACC	RMSE		ACC	RMSE
GA	BASIC	0.877 (0.835)	0.390		0.604 (0.550)	0.501
	EXTENDED	0.877 (0.830)	0.402		0.661 (0.622)	0.560
GP	BASIC	0.898 (0.823)	0.394		0.631 (0.580)	0.563
	EXTENDED	**0.888 (0.850)**	**0.351**		0.665 (0.633)	0.577
REG	BASIC	0.639	0.469		**0.71**	**0.432**

3.7 Regression Model

To obtain more accurate comparisons between our models and standard models in finance, we implemented Linear regression with absorbing indicators (as reported in [16, 5]), using the Stata software.

We accomplish this by estimating the following model 10 times, two versions for each PF:

$$\mathrm{y}_{ipt} = \beta_0 + \beta_1 \ln(1+(\mathrm{y}'_{pt-1}A_t)') + \gamma Controls + \delta_c + \delta_j + \delta_t + \varepsilon_{ipt}$$

Where the dependent variable y_{ipt} is a dummy variable that identifies firms that include pay feature P, during year t, in their CEO compensation contract. $\ln(1+(\mathrm{y}'_{pt-1}A_t)')$ is the logarithm of one plus the number of connected firms at time t that included the specific PF in their CEO compensation at time $t-1$. The number of connected firms that included pay feature P last year is the outcome of the product , where $(\mathbf{y}_{pt-1})'$ is a vector that identifies firms that include pay feature P during year $t-1$ and A'_t is the Adjacency matrix. We control for standard firm characteristics in our regressions. The specific controls we include are Log(Total assets), Market to book, Debt/TA and ROA (see Table 1). We also include compensation consultant fixed effects (δ_c), time fixed effects (δ_t), and industry fixed effects (δ_j) defined by four-digit SIC variable. The standard errors are clustered at the firm level.

In the last 5 tests (corresponding to the Adopt condition described above), we also modify our dependent variable to identify firms that newly adopt a specific PF in their CEO compensation.

Thus, the regression model's outputs include RMSE and ACC (see Table 3) for both the Predict and the Adopt conditions described in the previous section. Note, however, that this model can only be used to predict single PFs.

4 Results

Results are summarized in Tables 3–5. All tables show results from both GA and GP runs. Where applicable we compare our results to those obtained by the Regression model (see Section 3.7), marked as REG.

Table 3 shows our best results on the Option awards PF (single PF). The table contains both pay PF prediction and PF adoption results. Best test accuracy obtain is 0.85 (with a corresponding train accuracy of 0.888). Since these values are close, it is possible to infer that the best model in this category, namely GP with the Extended database, succeeded not only in predicting well, but also with generalizing effectively to test examples. **All success measures are high, and well above the Regression model's performance, which indicates a good result across the board**. The best GA predictors are not far behind, but still obtain somewhat lower scores.

Predictors evolved under the Adopt condition, attained less success with all measures. In this category, our regression model attained the best results. As stated above, such predictions are considerably more difficult for two reasons: (1) The number of relevant examples is small (2) PFs of the current year cannot be used for prediction (as its value is always zero in these conditions). Since next year's PFs are typically positively correlated with those of the current year (correlation values are approximately 0.7), this has a strong effect on prediction accuracy.

Also, GP predictors evolved with the Extended database performed better, plausibly due to utilizing information from more variables.

Table 4 Shows results evolving predictor which can simultaneously predict for all PF, under the Predict condition (as described in Section 3.6, data was insufficient to evolve Adopt results under these conditions).

GP still outperforms GA in these experiments attaining 86.4% accurate predictions on test examples (and 90% on training examples).

Table 4: All Pay-Features Predict Feature Table

Method	Params	Predict	
		ACC	RMSE
GA	BASIC	0.888 (0.831)	0.412
	EXTENDED	0.889 (0.832)	0.387
GP	**BASIC**	**0.900 (0.864)**	**0.369**
	EXTENDED	0.898 (0.856)	0.379

Table 5: Option and Absolute Pay-Features Adopt Feature Table

Method	Params	Adopt	
		ACC	RMSE
GA	BASIC	0.622 (0.540)	**0.500**
	EXTENDED	0.655 (0.600)	0.544
GP	BASIC	0.638 (0.578)	0.564
	EXTENDED	**0.640 (0.615)**	0.620

Surprisingly, performance did not degrade under these conditions, compared to predicting single PFs with "specialized" predictors. This result may be ascribed to the diversity of different examples, and show that our approach is scalable for this domain, finding better combinations with more variables.

Under these conditions as well, GP outperformed GA, and the Extended database proved more effective for training predictors. The reasons are probably similar to those stated above.

Table 5 Shows Adopt results evolving for Option and Absolute PFs together. The Adopt condition proved to be more difficult in these experiments as well, as reflected in all 3 measures of performance. Yet again, prediction made by GP individuals predictions for test examples were more accurate, compared to GA predictors, but the effect is not as salient as in the previous experiments.

Additionally, predictors evolved with the Extended database were still better than those evolved with the Basic database, demonstrating yet again that evolution has found ways to tap into the information contained in the relatively large group of variables. Still, overall prediction success in smaller, compared to the previous conditions.

5 Conclusions and Future Work

We evolved predictors for pay features in board networks, and important domain of empirical finance research, to which no EAs have been applied before. We evolved predictors using both genetic algorithms and genetic programming. We tested our predictors with groups of variables which are commonly used in this domain, under various experimental conditions. Our results show that our method yielded successful predictors, as quantified by several measures, including percentage of correct predictions and round mean squared error.

Our experiments provide 3 major results: (1) All predictors we evolved tackled this difficult domain successfully, yielding accurate predictions which attained an accuracy score of above 85% or above in several experiments. (2) GP outperformed GA when predicting for test examples under all experimental conditions, especially when using the Extended database, demonstrating that logical queries can lead to better predictions for pay features. (3) The best predictions were given by individuals which were evolved to predict multiple pay features with the *same set of weights*, yielding comparable (and even slightly better) performance on average than individuals evolved to predict specific pay features. Since we observed that all single PFs are equally difficult to predict, this is a rather surprising result.

We conclude that evolution (especially GP) has learned to reap the benefits of knowledge gained from being exposed to more diverse examples, later applying that knowledge to new examples successfully.

Individuals evolved in our experiments perform better than a regression model we constructed for this domain, using standard tools, and widely-used state-of-the-art methods. The potential of evolution to contribute to financial research is clear. Advantages of evolution come to the fore where regression models are limited: First, the applicability of our models to simultaneously predicting multiple pay features has been demonstrated in our experimental settings. Second, as stated above, regression models can only predict linear relation between variables, while GP can predict various types of inter-connections, including conditions (e.g. comparing variables), arbitrary mathematical functions of the input variables (which can be easily added), and more.

Possible continuations of our work include using different EA methods for this uncharted domain. For example, the problem of predicting multiple PFs can also be viewed as a multi-objective optimization problem [13], using a different fitness function for each PF, since typically each pay feature variable is predicted separately.

An additional novelty that evolution could bring to this domain is the ability to perform feature selection. The standard finance datasets mentioned in Section 3.1 contain thousands of variables, many of which are highly correlated–a major limitation for more traditional approaches. With evolution, it would be possible to examine connections between relatively large groups of variables easily, simply by plugging them into an EA model and evolving predictors for them.

And finally, our examinations show that evolution tends to hone in on a tree design early in the run, limiting its own search space to a degree. We believe that more population diversity may allow the evolutionary process to do even better. To that end we plan to incorporate diversity maintenance measures and have already begun to experiment with binning techniques similar to [23, 2]. Encouraged by our positive results, we have already begun to pursue some of the ideas presented above, and hope to report more results in the near future.

References

[1] Barnea, A. and Guedj, I. 2006. 'But, Mom, All the Other Kids Have One!'-CEO Compensation and Director Networks. In *1st Annual Conference on Empirical Legal Studies*. SSRN Electronic Journal.

[2] Benbassat, A. and Shafet, Y. 2017. A Simple Bucketing Based Approach to Diversity Maintenance. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion, GECCO '17*. ACM, New York, NY, USA, pp. 1559–1564.

[3] Bennett, B., Bettis, J. C., Gopalan, R., and Milbourn, T. 2017. Compensation goals and firm performance. *Journal of Financial Economics* 124, 2, pp. 307–330.

[4] Biethahn, J. and Nissen, V. 1994. Combinations of simulation and evolutionary algorithms in management science and economics. *Annals of Operations Research* 52, 4, pp. 181–208.

[5] Bizjak, J., Lemmon, M, and Whitby, R. 2009. Option backdating and board interlocks. *The Review of Financial Studies* 22, 11, pp. 4821–4847.

[6] Bouwman, C. H. S. 2011. Corporate governance propagation through overlapping directors. *The Review of Financial Studies* 24, 7, pp. 2358–2394.

[7] Cai, Y. and Sevilir, M. 2012. Board connections and M&A transactions. *Journal of Financial Economics* 103, 2, pp. 327–349.

[8] Dahya, J. and McConnell, J. J. 2005. Outside directors and corporate board decisions. *Journal of Corporate Finance* 11, 1–2, pp. 37–60.

[9] Dahya, J. and McConnell, J. J. 2007. Board composition, corporate performance, and the Cadbury committee recommendation. *Journal of Financial and Quantitative Analysis* 42, 3, pp. 535–564.

[10] Dahya, J., McConnell, J. J., and Travlos, N. G. 2002. The Cadbury committee, corporate performance, and top management turnover. *The Journal of Finance* 57, 1, pp. 461–483.

[11] Davis, G. F. 1991. Agents without principles? The spread of the poison pill through the intercorporate network. *Administrative science quarterly*, pp. 583–613.

[12] Davis, G. F. and Greve, H. R. 1997. Corporate elite networks and governance changes in the 1980s. *American journal of sociology* 103, 1, pp. 1–37.

[13] Deb, K. 2014. Multi-objective optimization. In *Search methodologies*. Springer, pp. 403–449.

[14] Elton, E. J., Gruber, M. J., and Blake, C. R. 2001. A first look at the accuracy of the CRSP mutual fund database and a comparison of the CRSP and Morningstar mutual fund databases. *The Journal of Finance* 56, 6, pp. 2415–2430.

[15] Fortin, F.-A., De Rainville, F.-M., Gardner, M.-A., Parizeau, M., and Gagné, C. 2012. DEAP: Evolutionary algorithms made easy. *Journal of Machine Learning Research* 13, Jul, pp. 2171–2175.

[16] Fracassi, C. and Tate, G. 2012. External networking and internal firm governance. *The Journal of Finance* 67, 1, pp. 153–194.

[17] Freeman, L. C. 1977. A set of measures of centrality based on betweenness. *Sociometry*, pp. 35–41.

[18] Gygax, A., Hazledine, M., and Martin, J. S. 2016. Are Fat Cats Copycats? Evidence from Director and Consultant Network Dynamics. In *2nd Annual Financial Institutions, Regulation and Corporate Governance Conference*. DOI: 10.2139/ssrn.2832531

[19] Hallock, K. F. 1997. Reciprocally interlocking boards of directors and executive compensation. *Journal of financial and Quantitative Analysis* 32, 3, pp. 331–344.

[20] Harrald, P. G. and Kamstra, M. 1997. Evolving artificial neural networks to combine financial forecasts. *IEEE Transactions on Evolutionary Computation* 1, 1, pp. 40–52.

[21] Horton, J., Millo, Y., and Serafeim, G. 2012. Resources or power? Implications of social networks on compensation and firm performance. *Journal of Business Finance & Accounting* 39, 3–4, pp. 399–426.

[22] Kantardzic, M. 2011. *Data mining: concepts, models, methods, and algorithms*. John Wiley & Sons.

[23] Martin, W. N., Lienig, J., and Cohoon, J. P. 1997. Island (Migration) models: evolutionary algorithms based on punctuated equilibria. In *Handbook of Evolutionary Computation (Eds: T. Baeck, D.B Fogel, Z. Michalewicz)*. Taylor & Francis.

[24] Murphy, K. P. 2012. *Machine Learning: a probabilistic perspective*. MIT Press.

[25] Nissen, V. 2018. Applications of Evolutionary Algorithms to Management Problems. In *Innovative Research Methodologies in Management*. Springer, pp. 211–235.

[26] Renneboog, L. and Zhao, Y. 2011. Us knows us in the UK: On director networks and CEO compensation. *Journal of Corporate Finance* 17, 4, pp. 1132–1157.

[27] Ulbricht, N. and Weiner, C. 2005. Worldscope meets Compustat: A comparison of financial databases. DOI: 10.2139/ssrn.871169

[28] Weisbach, M. S. 1998. Outside directors and CEO turnover. *Journal of financial Economics* 20, pp. 431–460.

[29] Wong, L. H. H., Gygax, A. F., and Wang, P. 2015. Board interlocking network and the design of executive compensation packages. *Social Networks* 41, pp. 85–100.