Introduction & Approach

Data:
~ sequence of tuples (t, q), where t relating an elapsed amount of time and to the quality q of the best solution discovered with t

Models Fitting Quality Φ: Φ(M) = 1/n \left(\sum_{i=1}^{n} (M(t) - q)^2 \right)q

Curve Fitting Method:
~ Levenberg-Marquardt algorithm
~ intelligent initialization strategy
~ multiple restarts
~ models:

<table>
<thead>
<tr>
<th>Name</th>
<th>Shortcut</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decay Model</td>
<td>DCM</td>
<td>A + B * exp(C \cdot t^p)</td>
</tr>
<tr>
<td>Logistic Model</td>
<td>LGM</td>
<td>A + B / (1 + exp(C * ln(t) + D))</td>
</tr>
<tr>
<td>Exp-Linear Model</td>
<td>ELM</td>
<td>A + B * exp(C * ln(t + D))</td>
</tr>
<tr>
<td>Gompertz Model</td>
<td>GPM</td>
<td>A + B / (1 + exp(C * exp(D * t)))</td>
</tr>
</tbody>
</table>

ANNs:
~ represent a much wider range of behavior
~ have significantly more parameters
~ the semantics of these parameters are too complex to be manually interpreted
~ might violate the constraint that any reasonable optimization process will never forget the best solution encountered
~ in our experiment we use feed-forward ANNs with a single hidden layer.

a. Generalized exponential decay (DCMP) y = M(t) = A + B * exp(C * t^p) for A ε [0, 50], B ε [50, 100], C ε [-0.1, -0.01], D ε [0.3, 3]
b. Generalized logistic model (LGMP) y = M(t) = A + B / (1 + exp(C * ln(t) + D)) for A ε [0, 50], B ε [50, 200], C ε [0.5, 0.8], D ε [0.3, 3]
c. Gompertz function (GMPM) y = M(t) = A + B / (1 + exp(C * exp(D * t))) for A ε [50, 100], B ε [50, 100], C ε [-3.2, -0.05], D ε [0.005, 0.15]
d. Exp-linear model (ELMP) y = M(t) = A + B * exp(C * ln(t + D)) for A ε [0, 50], B ε [50, 200], C ε [-0.1, -0.3], D ε [-0.9, 10]

Case Study: MAX-SAT

Problem Introduction & Experiment Setup:
~ six setups with hill climber algorithms: 3 search operator (1, 2, m-bit flip) and whether or not restarts are applied
~ 20 independent runs for each algorithm-instance combination
~ Maximum Satisfiability Problem (MAX-SAT) with n variables as case study

Model Parameters vs. Instance Features:
~ analysis relationship between model parameters and algorithm configuration
~ the smaller parameter A the better asymptotic results
~ parameter B has an upward trend for all algorithms with a rising number of variables/clauses of the MAX-SAT instances
~ parameter C and D both have approximately a negative linear association with the instance scale

Algorithm and Instance Classification:
~ samples feature: A, B, C, D and Φ
~ sample class: 6 in total, each algorithm setup as one class.
~ classifier: ANN with backpropagation, Support Vector Machine (SVM) with linear kernel, Gradient Boosting Decision Tree (GBDT)

Model Parameter Prediction:
~ the model parameters are clearly related to the instance features.
~ ANN directly predicts model parameters based on instance features.
~ first omit two instance scale parameters results and use ANN predict it and than compared them with the omitted true parameters by calculating the Φ
~ ANN wih linear activation function and grid search for weight decay
~ Root Mean Square Error (RMSE) for choosing optimal model

Prediction of Future Progress:
~ representing algorithm behavior as function can compute, for any point in time, which solution quality the algorithm likely has obtained
~ used for prediction future progress of algorithms in the running optimization process.
~ two methods, ANN and Curve Fitting
~ (train, test) = (50, 100) stands for predicting the complete algorithm behavior during the first 50 in the future based on the data points collecte