Archetype FAQ

Award-winning, patent-pending AI powered predictive modelling software.

What types of model does Archetype create?

Archetype produces two types of model.

- Binary – like application scores, behavioural scores for risk, marketing propensity: anything where you are trying to predict a true / false outcome or the probability of an event happening
- Continuous – where you are predicting the value of a piece of data, such as lifetime customer value, income, footfall, spend, expected loss

Archetype was designed with the credit risk market in mind, but it is completely data agnostic and can be used to generate models for any purpose where the above types of outcome can be predicted from the data that you hold.

How much data do I need to create an Archetype model?

As with all modelling tasks, the best models are created from bigger data sets. And even though it's AI-based, Archetype still needs enough data to work from. We'd recommend at least 10,000 rows of data with at least 1,000 examples of the outcome you are looking to predict. Your model development files can be up to 3Gb in size, with no limit on the number of data fields you can supply (although bigger models take longer to run). In general though, the more data the better. With Archetype, there are a number of techniques which help you to quickly identify the most predictive data inputs, so that you can get down to a manageable number of fields within the model: start with a sample of rows and use Archetype to choose the columns of data which show greatest promise, then filter down to a smaller number of data inputs and use all the rows data to generate the best model.

How does Archetype enable me to control the models I build?

Archetype offers considerable control over the creation of models, and incorporates a patent-pending algorithm which enables Archetype to produce models that adhere to business rules.

It enables the analyst to:

- Retain full control of the data that's considered for modelling
- Select and deselect candidate variables for inclusion in the model, based on a range of statistical measures, overriding Archetype's choice of variables
- Apply field-level constraints over each variable where required, defining how the model output should change with changes in each field value – for instance that credit risk should always increase as affordability worsens
- Categorise data values that should be regarded as equivalent, and rank order individual data values and groups of values
- Vary the topology of the underlying neural network – to add or remove neurons or neuron layers, change the number of iterations, control how much data contributes to each modelling iteration
- Store and retrieve project and data definitions (including bureau values) so that new models can be built without having to redefine model characteristics

Archetype FAQ

Award-winning, patent-pending
AI powered predictive modelling software.

How does Archetype enable me to explain the model outputs?

Archetype produces five different chart outputs which give a comprehensive explanation of how the model performs. This includes an overall performance chart, charts which look at the shape of the contribution of each variable to the overall output (expected versus actual), and charts which reveal the most influential variables within the model, which can be used as the basis of a monitoring report. As well as rank ordering the most influential variables, Archetype will also show you which variables aren't pulling their weight, and which non-included variables could add even more predictive power to the model. And it helps you avoid over-fitting by highlighting where the model isn't performing as it should.

Aren't the models really complex and time-consuming to run?

Creating a neural network-based model is a mathematically complicated process but, as a hosted solution, Archetype makes the process as easy as uploading a data set and clicking some buttons. The first pass of a new model can be created from scratch within a couple of hours; an existing model can be reconfigured and refreshed in as little as half an hour.

Once created, the resulting model code is lengthy but not complex. Archetype outputs model code in a range of different scripts which run efficiently in any modern decision system or data environment. Models can also be executed via an API, or run via Archetype's front end via a data upload.

How do Archetype models fit in to my company's governance policies?

Archetype's patent-pending algorithm enables you to determine up-front what rules the model should follow in respect of each individual data field. This level of control guarantees that the resulting model follows behaviours that you would currently demand within an existing scorecard.

Having set up a modelling project that adheres to your governance policy, future refreshes of the model automatically inherit all of the data rules and characteristics, meaning that you can replace a model more rapidly as its performance eventually deteriorates, safe in the knowledge that the rules that you have defined for the initial model still apply to the subsequent one.

What about model monitoring?

There is still a need to develop model monitoring for an Archetype model. Although the models are more complicated than a traditional score, the underlying process of monitoring variables is not, and existing monitoring templates can be used without needing large changes.

Additionally, the Archetype model can be simply re-run periodically, using refreshed data, to allow direct comparison between the existing model and a fully optimised one, supporting a highly informed decision on whether to process with a model update.

The charts that Archetype generates contribute to an assessment of how each variable contributes to the outcome, as well as the most important contributors to the model, which establishes a priority order for assessment of individual variables and a comparison with the model at the point of deployment.

Offline monitoring of the model variables can be developed to assess the impact of shifts in individual fields. Typically, there are more variables in an AI-based model than in a traditional scorecard. There are more variables to monitor, but the models tend to be more stable over a longer period of time.

Archetype FAQ

Award-winning, patent-pending AI powered predictive modelling software.

How do I access and deploy an Archetype model?

Archetype models can be built for you by the Jaywing team, or developed by your own team accessing the software – or a combination of both. The system is browser-based and hosted by Jaywing, with just a simple, secure data upload to undertake as a means of getting started with your modelling project.

Once the model is completed, the code for the model can be downloaded directly from the system in a range of scripts, for direct deployment within your decision system. Alternatively, a model API is available for you to call within your platform. The interface can also be used to score up a file of customers for batch use or testing.

How are models refreshed? Is Archetype a self-learning system?

Models can be re-configured as often as required using a refreshed data file. Because of the long-term nature of credit risk and similar outcome windows, Archetype is not designed to continually update the models it generates. Instead, control of the process resides with the expert credit analyst, who makes decisions about when a model is due for refresh, publishing a new version of the model for deployment whenever it is deemed necessary. Nevertheless, Archetype makes it very easy to refresh a model based on new data, so it is feasible to run much shorter update cycles than have been possible historically.

Refreshing a model is a quick process: upload the new data file, apply the existing settings from the initial model version, and click a button to rebuild. The process will take around 30 minutes.

How stable are Archetype models?

In our experience, Archetype models are more stable than traditional scorecards and degrade more slowly. This is because they generally include a higher number of data inputs which contribute more evenly to the overall outcome, with responsibility for performance dispersed across a larger number of variables, rather than being heavily reliant on a handful of them. The model training processes within Archetype are designed to make the model less reliant on any single variable.

Because shifts in individual variable distributions has a lower impact on the model as a whole, this means that the model typically retains its performance for longer, reducing the need for redevelopment or reconfiguration .

How powerful are Archetype models compared to an existing scorecard?

Each model is different, and the benefits can vary depending on the portfolio, the availability of data, the quality of the book and the purpose to which the model is put.

We have seen an uplift compared to a non-linear model in every comparison we have run, typically the improvement in model performance is in the region of 10% , but has ranged from 2.5% up to 19% with the main driver being the quantity of data available for modelling (the more data, the greater the uplift). This typically corresponds to a multi-million pound saving in bad debt when incorporated into credit strategy.

Archetype FAQ

Award-winning, patent-pending
AI powered predictive modelling software.

Doesn't applying constraints to Archetype models lose all the benefit?

Applying constraints, so that the model always behaves in a way that is acceptable to the business and to regulators, is an essential part of what Archetype does. It means that you can safely deploy an Archetype model, assured that it will not exhibit unwanted, non-explainable behaviour.

Comparing a fully constrained Archetype model with a neural network model which doesn't have any constraints typically sees only a minor reduction in performance – in the region of 0.5% on the gini uplift, and often much less. For models based on very noisy data, constraints can actually improve performance. As with the above explanation of the stability of the models, the neural network simply finds an alternative route to the best solution, meaning that constrained models are almost entirely as powerful as an unconstrained one.

How does Archetype avoid bias?

Bias can be avoided through the careful construction and use of data within the modelling step. It is a human consideration which can be enforced with the help of Archetype's controls. This would include ensuring that any fields which could introduce bias unwittingly are either excluded from the model, or modelled in such a way that bias is removed (for example, by grouping together gendered titles so that the model considers them as equivalents.

We are currently considering further system enhancements which will enable bias to be avoided as an automatic part of the modelling process.

What techniques are used within Archetype?

Archetype builds models using Neural Networks, which we chose because of their particular suitability for building credit scoring models.

We've developed our patent-pending training algorithms that can ensure that the models it generates adhere to business rules specified within the user interface – making the models controllable and explainable. All input fields can be individually constrained (or not, if a constraint is not relevant) such that they meet monotonicity requirements (for continuous data) and / or category ranking rules (for categorical data or specific data values within a continuous range).

Archetype does not include Machine Learning techniques other than Neural Networks, because (i) the results they generate would tend to be no more powerful than neural networks and (ii) the explainability problem has not been solved for other techniques.

How does Archetype change the role of credit analysts?

While Archetype generates well-performing models out of the box, there is a definite advantage in having expert users, who know the business's own data, to make sensible decisions about what data to include, how to exploit it for maximum value and how to best constrain the model to meet business requirements. So, credit analysts remain an important part of the model creation process.

Archetype removes a number of steps involved in the modelling process, vastly reduces the number of iterations needed to get to a final model, and makes it easier and quicker to revise and reconfigure models in future.

Archetype FAQ

Award-winning, patent-pending AI powered predictive modelling software.

Who should use Archetype?

Anyone seeking to build highly accurate predictive models from the data they hold. Archetype can be used by non-experts to build highly performant models, but is flexible and configurable enough for modelling experts to drive even higher performance, significantly improving on results obtainable via other approaches.

Is Archetype secure?

Very much so. Archetype is securely hosted by Jaywing; access is strictly controlled and is locked down to authorised access locations using IP address restrictions. Modelling projects can be shared by a user within your organisation to nominated users, but all data uploaded for modelling purposes is securely hosted and only accessible to your own analysts. All the outputs can be downloaded to your own network as needed. The Archetype software is used to generate the models but not to execute live scores or decisions.

When it comes to executing an Archetype model, you have a number of choices. The easiest implementation is within your own decision platform, under your control and security protocol. Alternatives include use of a model-specific API which can be executed securely from a range of hosting locations.