

Accelerating Model Deployment With Rook

A Case Study From The Trading Floor

Jean-Robert Avettand-Fenoel

jravettand at gmail.com

LondonR, Wednesday 7th September 2011

Disclaimer

Accelerating
Model
Deployment With
Rook

Jean-Robert
Avettand-Fenoel

Use Case

Possible
Alternatives

Practicalities

A New Model to
Deploy!

Is Rook For You?

References

Disclaimer

The financial techniques and models presented here have been developed and are only used to illustrate the R techniques I demonstrate.

Their validity hasn't been tested at all and they shouldn't be taken as an investment advice.

Agenda

Accelerating
Model
Deployment With
Rook

Jean-Robert
Avettand-Fenoel

Use Case

Possible
Alternatives

Practicalities

A New Model to
Deploy!

Is Rook For You?

References

1 Use Case

2 Possible Alternatives

3 Practicalities

4 A New Model to Deploy!

5 Is Rook For You?

Use Case

Accelerating
Model
Deployment With
Rook

Jean-Robert
Avettand-Fenoel

Use Case

Possible
Alternatives

Practicalities

A New Model to
Deploy!

Is Rook For You?

References

Use Case

Accelerating
Model
Deployment With
Rook

Jean-Robert
Avettand-Fenoel

Use Case

Possible
Alternatives

Practicalities

A New Model to
Deploy!

Is Rook For You?

References

- The traders ask the strategists to help them predict the close of an index.

Use Case

Accelerating
Model
Deployment With
Rook

Jean-Robert
Avettand-Fenoel

Use Case

Possible
Alternatives

Practicalities

A New Model to
Deploy!

Is Rook For You?

References

- The traders ask the strategists to help them predict the close of an index.
- As the day goes on, the value of the index changes. The better the traders' knowledge of where the index might close on the day, the better their ability to manage their book.

Use Case

Accelerating
Model
Deployment With
Rook

Jean-Robert
Avettand-Fenoel

Use Case

Possible
Alternatives

Practicalities

A New Model to
Deploy!

Is Rook For You?

References

- The traders ask the strategists to help them predict the close of an index.
- As the day goes on, the value of the index changes. The better the traders' knowledge of where the index might close on the day, the better their ability to manage their book.
- The strategists provide the traders a system to help them. The system can be queried at any time during the day and give an answer as to where the index might close on the day.

Use Case

Accelerating
Model
Deployment With
Rook

Jean-Robert
Avettand-Fenoel

Use Case

Possible
Alternatives

Practicalities

A New Model to
Deploy!

Is Rook For You?

References

- The traders ask the strategists to help them predict the close of an index.
- As the day goes on, the value of the index changes. The better the traders' knowledge of where the index might close on the day, the better their ability to manage their book.
- The strategists provide the traders a system to help them. The system can be queried at any time during the day and give an answer as to where the index might close on the day.

Problem

Assuming that we have a model to predict the close, how do we implement the system and make it easy to update?

Possible Alternatives

Accelerating
Model
Deployment With
Rook

Jean-Robert
Avettand-Fenoel

Use Case

Possible
Alternatives

Practicalities

A New Model to
Deploy!

Is Rook For You?

References

Possible Alternatives

Alternative 1

Implementation Mathematical Software to C++ to VBA to Excel

Pros Excel well spread

Cons Hard to translate to C++ and VBA, version control of Excel spreadsheet is a problem

Time To Deploy After development, from days to weeks

Accelerating
Model
Deployment With
Rook

Jean-Robert
Avettand-Fenoel

Use Case

Possible
Alternatives

Practicalities

A New Model to
Deploy!

Is Rook For You?

References

Possible Alternatives

Alternative 1

Implementation Mathematical Software to C++ to VBA to Excel

Pros Excel well spread

Cons Hard to translate to C++ and VBA, version control of Excel spreadsheet is a problem

Time To Deploy After development, from days to weeks

Alternative 2

Implementation R to RExcel to Excel

Pros Excel well spread, easy to link R and RExcel

Cons Need R and RExcel on all machines, problem of version control

Time To Deploy After development, a matter of days

Accelerating
Model
Deployment With
Rook

Jean-Robert
Avettand-Fenoel

Use Case

Possible
Alternatives

Practicalities

A New Model to
Deploy!

Is Rook For You?

References

Possible Alternatives

Alternative 1

Implementation Mathematical Software to C++ to VBA to Excel

Pros Excel well spread

Cons Hard to translate to C++ and VBA, version control of Excel spreadsheet is a problem

Time To Deploy After development, from days to weeks

Alternative 2

Implementation R to RExcel to Excel

Pros Excel well spread, easy to link R and RExcel

Cons Need R and RExcel on all machines, problem of version control

Time To Deploy After development, a matter of days

Alternative 3

Implementation R to Browser via **Rook**

Pros Access via web browser, no version control for users

Cons Less convenient to access/manipulate data

Time To Deploy After development, a matter of hours

Practicalities

Accelerating
Model
Deployment With
Rook

Jean-Robert
Avettand-Fenoel

Use Case

Possible
Alternatives

Practicalities

A New Model to
Deploy!

Is Rook For You?

References

We will follow these steps

- 1 Introduce the prediction model
- 2 Build it into R
- 3 Create an output of the model
- 4 Transfer the output to Rook

k-Nearest Neighbor to the Rescue

Accelerating
Model
Deployment With
Rook

Jean-Robert
Avettand-Fenoel

Use Case

Possible
Alternatives

Practicalities

A New Model to
Deploy!

Is Rook For You?

References

k-Nearest Neighbor to the Rescue

Accelerating
Model
Deployment With
Rook

Jean-Robert
Avettand-Fenoel

Use Case

Possible
Alternatives

Practicalities

A New Model to
Deploy!

Is Rook For You?

References

- We use the k-nearest neighbor algorithm to predict the close. (*"amongst the simplest of all machine learning algorithms"*, Wikipedia)

k-Nearest Neighbor to the Rescue

- We use the k-nearest neighbor algorithm to predict the close. (*"amongst the simplest of all machine learning algorithms"*, Wikipedia)
- Given today's index behaviour from 8am to, say, 2pm, which days in the past (the neighbors) behaved similarly (the nearest)?

Accelerating

Model

Deployment With
Rook

Jean-Robert

Avettand-Fenoel

Use Case

Possible
Alternatives

Practicalities

A New Model to
Deploy!

Is Rook For You?

References

k-Nearest Neighbor to the Rescue

- We use the k-nearest neighbor algorithm to predict the close. (*"amongst the simplest of all machine learning algorithms"*, Wikipedia)
- Given today's index behaviour from 8am to, say, 2pm, which days in the past (the neighbors) behaved similarly (the nearest)?
- Once we know these days, we can extrapolate where the index might end the day.

Accelerating

Model

Deployment With
Rook

Jean-Robert

Avettand-Fenoel

Use Case

Possible
Alternatives

Practicalities

A New Model to
Deploy!

Is Rook For You?

References

k-Nearest Neighbor to the Rescue

- We use the k-nearest neighbor algorithm to predict the close. (*"amongst the simplest of all machine learning algorithms"*, Wikipedia)
- Given today's index behaviour from 8am to, say, 2pm, which days in the past (the neighbors) behaved similarly (the nearest)?
- Once we know these days, we can extrapolate where the index might end the day.
- How do we define the nearest? Use a distance, for instance, the Euclidean distance.

Accelerating

Model

Deployment With
Rook

Jean-Robert

Avettand-Fenoel

Use Case

Possible
Alternatives

Practicalities

A New Model to
Deploy!

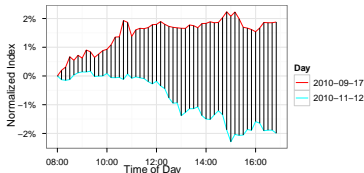
Is Rook For You?

References

k-Nearest Neighbor to the Rescue

- We use the k-nearest neighbor algorithm to predict the close. (*"amongst the simplest of all machine learning algorithms"*, Wikipedia)
- Given today's index behaviour from 8am to, say, 2pm, which days in the past (the neighbors) behaved similarly (the nearest)?
- Once we know these days, we can extrapolate where the index might end the day.
- How do we define the nearest? Use a distance, for instance, the Euclidean distance.

$$d_E(X, Y) = \sqrt{\sum_{t=t_0}^{t_N} (X_t - Y_t)^2}$$



The Prediction Function

Accelerating
Model
Deployment With
Rook

Jean-Robert
Avettand-Fenoel

Use Case

Possible
Alternatives

Practicalities

A New Model to
Deploy!

Is Rook For You?

References

The Prediction Function

Accelerating
Model
Deployment With
Rook

Jean-Robert
Avettand-Fenoel

Use Case

Possible
Alternatives

Practicalities

A New Model to
Deploy!

Is Rook For You?

References

Assuming we have some well conditioned data, `distMethod` is our generic distance method which takes the time series as an input:

The Prediction Function

Accelerating
Model
Deployment With
Rook

Jean-Robert
Avetrand-Fenoel

Use Case

Possible
Alternatives

Practicalities

A New Model to
Deploy!

Is Rook For You?

References

Assuming we have some well conditioned data, `distMethod` is our generic distance method which takes the time series as an input:

```
> predictClose <- function(oldDays, newDay, distMethod, k=10) {  
+  
+   # aggregate the past and the present  
+   aggDays <- cbind(oldDays[1:nrow(newDay)],, newDay)  
+  
+   # compute the distance between days  
+   distDays <- as.matrix(distMethod(aggDays))  
+  
+   # take the k nearest neighbours  
+   neighbs <- which(rank(distDays[,ncol(aggDays)],-ncol(aggDays))<=k)  
+  
+   # output that on a chart (only ggplot2 code)  
+   makePlot(oldDays, newDay, neighbs, k)  
+ }  
> # run the function with the Euclidean distance  
> predictClose(oldDays, newDay, distMethod=euclideanDist, k=10)
```

The Prediction Function

Accelerating
Model
Deployment With
Rook

Jean-Robert
Avettand-Fenoel

Use Case

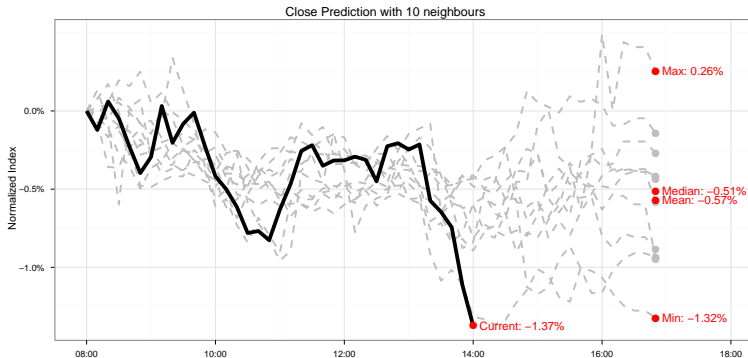
Possible
Alternatives

Practicalities

A New Model to
Deploy!

Is Rook For You?

References



Conversion to Rook - Step 1

Accelerating
Model
Deployment With
Rook

Jean-Robert
Avettand-Fenoel

Use Case

Possible
Alternatives

Practicalities

A New Model to
Deploy!

Is Rook For You?

References

Now comes the interesting part where the model has been developed and we make it available via Rook.

Conversion to Rook - Step 1

Accelerating
Model
Deployment With
Rook

Jean-Robert
Avettand-Fenoel

Use Case

Possible
Alternatives

Practicalities

A New Model to
Deploy!

Is Rook For You?

References

Now comes the interesting part where the model has been developed and we make it available via Rook.

1 Load Rook and start the web server

```
> require(Rook)
> s <- Rhttpd$new()
> s$start(listen='127.0.0.1',quiet=T)
```

Conversion to Rook - Step 2

Accelerating
Model
Deployment With
Rook

Jean-Robert
Avettand-Fenoel

Use Case

Possible
Alternatives

Practicalities

A New Model to
Deploy!

Is Rook For You?

References

2 Create an app for the web server (the biggest part)

```
> ClosePredictApp <- Builder$new(  
+  
+   # map specific behaviors according to URL  
+   URLMap$new('~/.*\\.html$' = function(env){  
+     req <- Request$new(env); res <- Response$new()  
+  
+     # add an image in html, which will be produced below  
+     res$write(paste('', sep=''))  
+     res$finish() },  
+  
+   # generate image on the fly when URL is a .png  
+   '~/.*\\.png$' = function(env){  
+     req <- Request$new(env); res <- Response$new()  
+     res$header('Content-type', 'image/png')  
+     t <- tempfile(); png(file=t,width=1040,height=585)  
+  
+     # update the data and use the very same function as before  
+     newDay <- updateIndexData()  
+     predictClose(oldDays, newDay, distMethod=euclideanDist, k=10)  
+  
+     dev.off(); res$body <- t; names(res$body) <- 'file'  
+     res$finish()  
+     }, '~.*' = Redirect$new('/index.html'))  
+ )
```

Conversion to Rook - Step 3

3 Add the app to the server and browse!

```
> s$add(app=ClosePredictApp, name='ClosePredict')  
> s$print()
```

```
Server started on 127.0.0.1:13043
```

```
[1] RookTest      http://127.0.0.1:13043/custom/RookTest  
[2] ClosePredict  http://127.0.0.1:13043/custom/ClosePredict
```

Call `browse()` with an index number or name to run an application.

```
> s$browse('ClosePredict')  
> # then when you're done  
> s$stop()
```

Conversion to Rook - Step 3

Accelerating
Model
Deployment With
Rook

Jean-Robert
Avettand-Fenoel

Use Case

Possible
Alternatives

Practicalities

A New Model to
Deploy!

Is Rook For You?

References

3 Add the app to the server and browse!

```
> s$add(app=ClosePredictApp, name='ClosePredict')  
> s$print()
```

```
Server started on 127.0.0.1:13043
```

```
[1] RookTest      http://127.0.0.1:13043/custom/RookTest
```

```
[2] ClosePredict http://127.0.0.1:13043/custom/ClosePredict
```

Call `browse()` with an index number or name to run an application.

```
> s$browse('ClosePredict')  
> # then when you're done  
> s$stop()
```

Change the IP address to your own address, and you can send the URL to other users

Conversion to Rook

Accelerating
Model
Deployment With
Rook

Jean-Robert
Avettand-Fenoel

Use Case

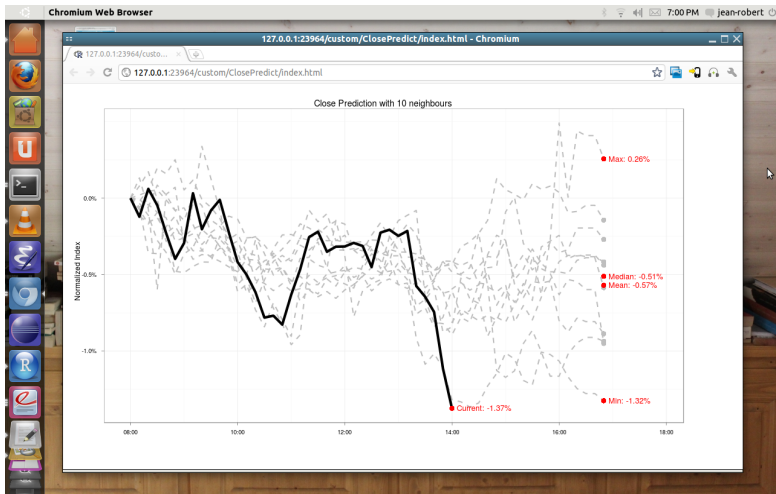
Possible
Alternatives

Practicalities

A New Model to
Deploy!

Is Rook For You?

References



Enhancements to the Model

Accelerating
Model
Deployment With
Rook

Jean-Robert
Avettand-Fenoel

Use Case

Possible
Alternatives

Practicalities

**A New Model to
Deploy!**

Is Rook For You?

References

Enhancements to the Model

Maybe the Euclidean distance was not the best distance...
What if we look at the Markov Operator distance De Gregorio and Maria Iacus (2010)?
Available as the `MOdist` function in the `sde` package.

Accelerating
Model

Deployment With
Rook

Jean-Robert
Avetrand-Fenoel

Use Case

Possible
Alternatives

Practicalities

A New Model to
Deploy!

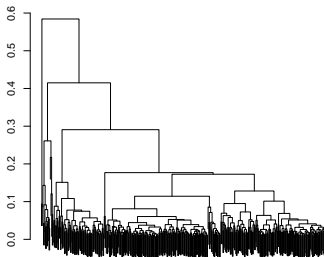
Is Rook For You?

References

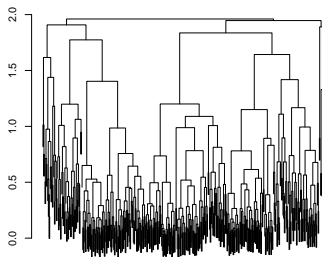
Enhancements to the Model

Maybe the Euclidean distance was not the best distance...
What if we look at the Markov Operator distance De Gregorio and Maria Iacus (2010)?
Available as the `MOdist` function in the `sde` package.
If we compare the clustering abilities of the two measures:

Clustering with Euclidean Distance

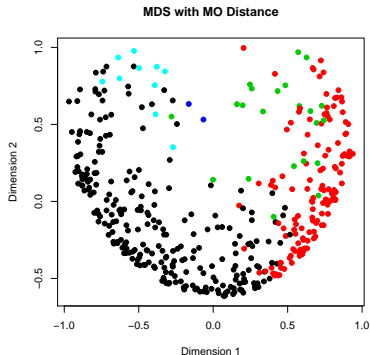
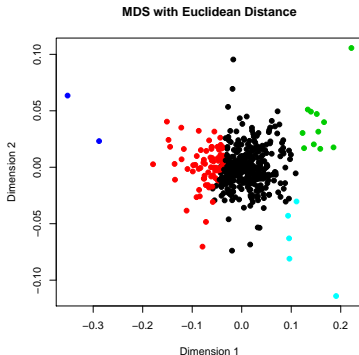


Clustering with MO Distance



Enhancements to the Model

Maybe the Euclidean distance was not the best distance...
What if we look at the Markov Operator distance De Gregorio and Maria Iacus (2010)?
Available as the `MOdist` function in the `sde` package.
If we compare the clustering abilities of the two measures:



How to Deploy the New Model

Accelerating
Model
Deployment With
Rook

Jean-Robert
Avettand-Fenoel

Use Case

Possible
Alternatives

Practicalities

**A New Model to
Deploy!**

Is Rook For You?

References

How to Deploy the New Model

Accelerating
Model
Deployment With
Rook

Jean-Robert
Avettand-Fenoel

Use Case

Possible
Alternatives

Practicalities

A New Model to
Deploy!

Is Rook For You?

References

In your development framework, this is done in one line:

```
> predictClose(oldDays, newDay, distMethod=MOdist, k=10)
```

How to Deploy the New Model

Accelerating
Model

Deployment With
Rook

Jean-Robert
Avettand-Fenoel

Use Case

Possible
Alternatives

Practicalities

A New Model to
Deploy!

Is Rook For You?

References

In your development framework, this is done in one line:

```
> predictClose(oldDays, newDay, distMethod=MOdist, k=10)
```

To deploy your new model, this is one line as well: the same one!

How to Deploy the New Model

Accelerating
Model

Deployment With
Rook

Jean-Robert
Avettand-Fenoel

Use Case

Possible
Alternatives

Practicalities

A New Model to
Deploy!

Is Rook For You?

References

In your development framework, this is done in one line:

```
> predictClose(oldDays, newDay, distMethod=M0dist, k=10)
```

To deploy your new model, this is one line as well: the same one!

- Instead of recoding M0dist in C++, updating the spreadsheet, and making sure everyone has the correct version (Alternative 1)

How to Deploy the New Model

Accelerating
Model
Deployment With
Rook

Jean-Robert
Avettand-Fenoel

Use Case

Possible
Alternatives

Practicalities

A New Model to
Deploy!

Is Rook For You?

References

In your development framework, this is done in one line:

```
> predictClose(oldDays, newDay, distMethod=MOdist, k=10)
```

To deploy your new model, this is one line as well: the same one!

- Instead of recoding MOdist in C++, updating the spreadsheet, and making sure everyone has the correct version (Alternative 1)
- Instead of updating the RExcel spreadsheet, and making sure everyone has the correct version (and everyone downloaded the new package) (Alternative 2)

How to Deploy the New Model

Accelerating
Model
Deployment With
Rook

Jean-Robert
Avettand-Fenoel

Use Case

Possible
Alternatives

Practicalities

A New Model to
Deploy!

Is Rook For You?

References

In your development framework, this is done in one line:

```
> predictClose(oldDays, newDay, distMethod=MOdist, k=10)
```

To deploy your new model, this is one line as well: the same one!

- Instead of recoding MOdist in C++, updating the spreadsheet, and making sure everyone has the correct version (Alternative 1)
- Instead of updating the RExcel spreadsheet, and making sure everyone has the correct version (and everyone downloaded the new package) (Alternative 2)
- You can even embed additional features with a bit of HTML, like the choice of the distance, the choice of the number of neighbours,...

How to Deploy the New Model

Accelerating
Model
Deployment With
Rook

Jean-Robert
Avettand-Fenoel

Use Case

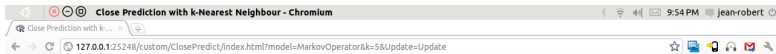
Possible
Alternatives

Practicalities

A New Model to
Deploy!

Is Rook For You?

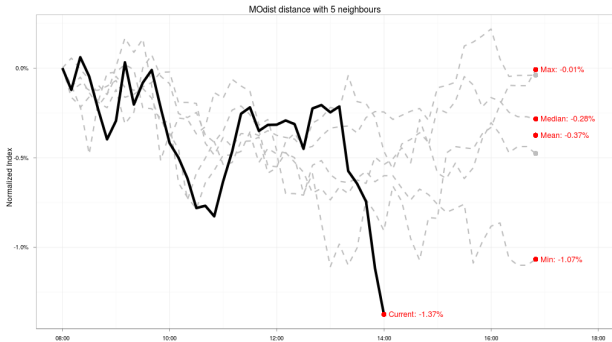
References



Close Prediction with k-Nearest Neighbour

Model type:

Number of neighbours:



Pros and Cons of Rook

Accelerating
Model
Deployment With
Rook

Jean-Robert
Avettand-Fenoel

Use Case

Possible
Alternatives

Practicalities

A New Model to
Deploy!

Is Rook For You?

References

Pros and Cons of Rook

Accelerating
Model
Deployment With
Rook

Jean-Robert
Avettand-Fenoel

Use Case

Possible
Alternatives

Practicalities

A New Model to
Deploy!

Is Rook For You?

References

Rook is for you if:

- You want to transfer your model output easily to a web interface
- The end user is not technical/not a programmer
- You need to reduce the database/CPU workload (Rook transfers it from the client to the server)

Pros and Cons of Rook

Accelerating
Model

Deployment With
Rook

Jean-Robert
Avettand-Fenoel

Use Case

Possible
Alternatives

Practicalities

A New Model to
Deploy!

Is Rook For You?

References

Rook is for you if:

- You want to transfer your model output easily to a web interface
- The end user is not technical/not a programmer
- You need to reduce the database/CPU workload (Rook transfers it from the client to the server)

Rook is maybe not for you if:

- You don't like HTML/web programming
- The end user needs to directly manipulate the data (possibly solved with custom csv file generation...)

Possible Use and Improvements

Accelerating
Model
Deployment With
Rook

Jean-Robert
Avettand-Fenoel

Use Case

Possible
Alternatives

Practicalities

A New Model to
Deploy!

Is Rook For You?

References

Possible Use and Improvements

Possible use:

- Continuous risk monitoring for portfolio managers
- Intraday signals for pairs trading
- Also not in finance: easy generation of any on-the-fly chart/statistics within one R session

Accelerating
Model

Deployment With
Rook

Jean-Robert
Avettand-Fenoel

Use Case

Possible
Alternatives

Practicalities

A New Model to
Deploy!

Is Rook For You?

References

Possible Use and Improvements

Possible use:

- Continuous risk monitoring for portfolio managers
- Intraday signals for pairs trading
- Also not in finance: easy generation of any on-the-fly chart/statistics within one R session

Improvements:

- Offload some of the work to an other R session via Rserve
- Inclusion in bigger websites with some data intensive contents
- More complex web content with JavaScript and/or HTML5

Possible Use and Improvements

Accelerating
Model
Deployment With
Rook

Jean-Robert
Avetrand-Fenoel

Use Case

Possible
Alternatives

Practicalities

A New Model to
Deploy!

Is Rook For You?

References

Possible use:

- Continuous risk monitoring for portfolio managers
- Intraday signals for pairs trading
- Also not in finance: easy generation of any on-the-fly chart/statistics within one R session

Improvements:

- Offload some of the work to an other R session via Rserve
- Inclusion in bigger websites with some data intensive contents
- More complex web content with JavaScript and/or HTML5

Also available:

- RApache (<http://www.rapache.net/>)
- openCPU (<http://www.opencpu.org/>)
- googleVis (<http://code.google.com/p/google-motion-charts-with-r/>)

References

Accelerating
Model
Deployment With
Rook

Jean-Robert
Avettand-Fenoel

Use Case

Possible
Alternatives

Practicalities

A New Model to
Deploy!

Is Rook For You?

References

De Gregorio, A. and Maria Iacus, S. (2010). Clustering of discretely observed diffusion processes. *Computational Statistics & Data Analysis*, 54(2):598–606.

Horner, J. (2011). *Rook: Rook - a web server interface for R*. R package version 1.0-2.

Iacus, S. M. (2009). *sde: Simulation and Inference for Stochastic Differential Equations*. R package version 2.0.10.