Ensembles and Combining Models

Christopher Cooksey, FCAS, MAAA
Head Actuary, Data & Analytics

28 March 2017

GUIDEWIRE

Agenda

- Rationale and Effectiveness of Ensembles
- Basic Approaches – Bagging and Boosting
- Complexity – Issues and Advantages
- Combining Linear Regression and Ensembles

Rationale and Effectiveness of Ensembles

What is the “best” model?

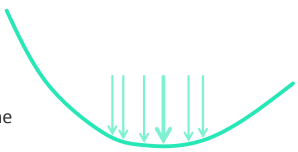
There isn’t only one correct model.

Consider credibility-weighting a statewide average with a countrywide average.

What is the “best” model?

If you have two models, each of which perform similarly from a statistical perspective, which do you choose?

Normally we work with some function to define “best.”



Multiplicity of Models

“…there is often a multitude of different descriptions [equations f(x)] in a class of functions giving about the same minimum error rate.”

Breiman, L. (2001). Statistical Modeling: The Two Cultures. *Statistical Science*, Vol. 16, No. 3.

“Data will often point with almost equal emphasis on several possible models, and it is important that the statistician recognize and accept this.”

McCullagh, P. and Nelder, J. (1989). *Generalized Linear Models*.

AN UNREALISTIC ILLUSTRATION

Ground Rules

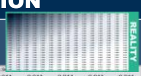
1) We get to know reality & compare our models directly.
2) Assume the numbers are frequency relativities.
3) Volume is limited; we can only divide the data into three equally-sized groups.
4) Model predictions are just the average for each defined group.

AN UNREALISTIC ILLUSTRATION

[illegible]	[illegible]	1.801	1.732	1.665	1.539	1.480	1.423	1.316	1.265	1.217	1.125	1.082	1.040	1.000
[illegible]	1.873	1.732	1.665	1.601	1.480	1.423	1.369	1.265	1.217	1.170	1.082	1.040	1.000	1.000
1.873	1.801	1.665	1.601	1.539	1.423	1.369	1.316	1.217	1.170	1.125	1.040	1.000	1.000	1.000
1.801	1.732	1.601	1.539	1.480	1.369	1.316	1.265	1.170	1.125	1.082	1.000	1.000	1.000	1.000
1.732	1.665	1.539	1.480	1.423	1.316	1.265	1.217	1.125	1.082	1.040	1.000	1.000	1.000	1.000
1.665	1.601	1.480	1.423	1.369	1.265	1.217	1.170	1.082	1.040	1.000	1.000	1.000	1.000	1.000
1.601	1.539	1.423	1.369	1.316	1.217	1.170	1.125	1.040	1.000	1.000	1.000	1.000	1.000	0.980
1.539	1.480	1.369	1.316	1.265	1.170	1.125	1.082	1.000	1.000	1.000	1.000	1.000	1.000	0.980
1.480	1.423	1.316	1.265	1.217	1.125	1.082	1.040	1.000	1.000	1.000	1.000	1.000	0.980	0.960
1.423	1.369	1.265	1.217	1.170	1.082	1.040	1.000	1.000	1.000	1.000	1.000	1.000	0.980	0.960
1.369	1.316	1.217	1.170	1.125	1.040	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.980	0.960
1.316	1.265	1.170	1.125	1.082	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.980	0.960	0.941
1.265	1.217	1.125	1.082	1.040	1.000	1.000	1.000	1.000	1.000	1.000	0.980	0.960	0.941	0.922
1.217	1.170	1.082	1.040	1.000	1.000	1.000	1.000	1.000	1.000	0.980	0.960	0.941	0.922	0.904
1.170	1.125	1.040	1.000	1.000	1.000	1.000	1.000	1.000	0.980	0.960	0.941	0.922	0.904	0.886
1.125	1.082	1.000	1.000	1.000	1.000	1.000	1.000	0.980	0.960	0.941	0.922	0.904	0.886	0.868
1.082	1.040	1.000	1.000	1.000	1.000	1.000	1.000	0.980	0.960	0.941	0.922	0.904	0.886	0.868
1.040	1.000	1.000	1.000	1.000	1.000	1.000	0.980	0.960	0.941	0.922	0.904	0.886	0.868	0.851
1.000	1.000	1.000	1.000	1.000	1.000	0.980	0.960	0.941	0.922	0.904	0.886	0.868	0.851	0.834
1.000	1.000	1.000	1.000	1.000	0.980	0.960	0.941	0.922	0.904	0.886	0.868	0.851	0.834	0.817
1.000	1.000	1.000	1.000	0.980	0.960	0.941	0.922	0.904	0.886	0.868	0.851	0.834	0.817	0.801
1.000	1.000	1.000	1.000	0.980	0.960	0.941	0.922	0.904	0.886	0.868	0.851	0.834	0.817	0.801
1.000	1.000	1.000	0.980	0.960	0.941	0.922	0.904	0.886	0.868	0.851	0.834	0.817	0.801	0.785
1.000	1.000	0.980	0.960	0.941	0.922	0.904	0.886	0.868	0.851	0.834	0.817	0.801	0.785	0.769

REALITY

AN UNREALISTIC ILLUSTRATION



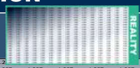
MODEL 1

Group relatively homogeneous business together.

Sum of the squared error = 13.48

1.260	1.260	1.260	1.260	1.260	1.059	1.059	1.059	1.059	1.059	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
1.260	1.260	1.260	1.260	1.260	1.059	1.059	1.059	1.059	1.059	0.944	0.944	0.944	0.944	0.944
1.260	1.260	1.260	1.260	1.260	1.059	1.059	1.059	1.059	1.059	0.944	0.944	0.944	0.944	0.944
1.260	1.260	1.260	1.260	1.260	1.059	1.059	1.059	1.059	1.059	0.944	0.944	0.944	0.944	0.944
1.260	1.260	1.260	1.260	1.260	1.059	1.059	1.059	1.059	1.059	0.944	0.944	0.944	0.944	0.944
1.260	1.260	1.260	1.260	1.260	1.059	1.059	1.059	1.059	1.059	0.944	0.944	0.944	0.944	0.944
1.260	1.260	1.260	1.260	1.260	1.059	1.059	1.059	1.059	1.059	0.944	0.944	0.944	0.944	0.944
1.260	1.260	1.260	1.260	1.260	1.059	1.059	1.059	1.059	1.059	0.944	0.944	0.944	0.944	0.944
1.260	1.260	1.260	1.260	1.260	1.059	1.059	1.059	1.059	1.059	0.944	0.944	0.944	0.944	0.944
1.260	1.260	1.260	1.260	1.260	1.059	1.059	1.059	1.059	1.059	0.944	0.944	0.944	0.944	0.944
1.260	1.260	1.260	1.260	1.260	1.059	1.059	1.059	1.059	1.059	0.944	0.944	0.944	0.944	0.944
1.260	1.260	1.260	1.260	1.260	1.059	1.059	1.059	1.059	1.059	0.944	0.944	0.944	0.944	0.944
1.260	1.260	1.260	1.260	1.260	1.059	1.059	1.059	1.059	1.059	0.944	0.944	0.944	0.944	0.944
1.260	1.260	1.260	1.260	1.260	1.059	1.059	1.059	1.059	1.059	0.944	0.944	0.944	0.944	0.944
1.260	1.260	1.260	1.260	1.260	1.059	1.059	1.059	1.059	1.059	0.944	0.944	0.944	0.944	0.944
1.260	1.260	1.260	1.260	1.260	1.059	1.059	1.059	1.059	1.059	0.944	0.944	0.944	0.944	0.944
1.260	1.260	1.260	1.260	1.260	1.059	1.059	1.059	1.059	1.059	0.944	0.944	0.944	0.944	0.944
1.260	1.260	1.260	1.260	1.260	1.059	1.059	1.059	1.059	1.059	0.944	0.944	0.944	0.944	0.944
1.260	1.260	1.260	1.260	1.260	1.059	1.059	1.059	1.059	1.059	0.944	0.944	0.944	0.944	0.944
1.260	1.260	1.260	1.260	1.260	1.059	1.059	1.059	1.059	1.059	0.944	0.944	0.944	0.944	0.944
1.260	1.260	1.260	1.260	1.260	1.059	1.059	1.059	1.059	1.059	0.944	0.944	0.944	0.944	0.944
1.260	1.260	1.260	1.260	1.260	1.059	1.059	1.059	1.059	1.059	0.944	0.944	0.944	0.944	0.944
1.260	1.260	1.260	1.260	1.260	1.059	1.059	1.059	1.059	1.059	0.944	0.944	0.944	0.944	0.944
1.260	1.260	1.260	1.260	1.260	1.059	1.059	1.059	1.059	1.059	0.944	0.944	0.944	0.944	0.944

AN UNREALISTIC ILLUSTRATION



MODEL 2

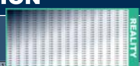
A different way of splitting the data.

Sum of the squared error = 11.63

1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287
1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287
1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287
1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287
1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287
1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287
1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287	1.287
1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048
1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048
1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048
1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048
1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048
1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048
1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048
1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048	1.048
0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929
0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929
0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929
0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929
0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929
0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929
0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929
0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929	0.929

GUIDEWIRE Page 10

AN UNREALISTIC ILLUSTRATION



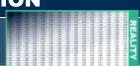
ENSEMBLE Models 1 & 2

Sum of the squared error = 9.02

1.274	1.274	1.274	1.274	1.274	1.173	1.173	1.173	1.173	1.173	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
1.274	1.274	1.274	1.274	1.274	1.173	1.173	1.173	1.173	1.173	1.116	1.116	1.116	1.116	1.116
1.274	1.274	1.274	1.274	1.274	1.173	1.173	1.173	1.173	1.173	1.116	1.116	1.116	1.116	1.116
1.274	1.274	1.274	1.274	1.274	1.173	1.173	1.173	1.173	1.173	1.116	1.116	1.116	1.116	1.116
1.274	1.274	1.274	1.274	1.274	1.173	1.173	1.173	1.173	1.173	1.116	1.116	1.116	1.116	1.116
1.274	1.274	1.274	1.274	1.274	1.173	1.173	1.173	1.173	1.173	1.116	1.116	1.116	1.116	1.116
1.274	1.274	1.274	1.274	1.274	1.173	1.173	1.173	1.173	1.173	1.116	1.116	1.116	1.116	1.116
1.274	1.274	1.274	1.274	1.274	1.173	1.173	1.173	1.173	1.173	1.116	1.116	1.116	1.116	1.116
1.154	1.154	1.154	1.154	1.154	1.053	1.053	1.053	1.053	1.053	0.996	0.996	0.996	0.996	0.996
1.154	1.154	1.154	1.154	1.154	1.053	1.053	1.053	1.053	1.053	0.996	0.996	0.996	0.996	0.996
1.154	1.154	1.154	1.154	1.154	1.053	1.053	1.053	1.053	1.053	0.996	0.996	0.996	0.996	0.996
1.154	1.154	1.154	1.154	1.154	1.053	1.053	1.053	1.053	1.053	0.996	0.996	0.996	0.996	0.996
1.154	1.154	1.154	1.154	1.154	1.053	1.053	1.053	1.053	1.053	0.996	0.996	0.996	0.996	0.996
1.154	1.154	1.154	1.154	1.154	1.053	1.053	1.053	1.053	1.053	0.996	0.996	0.996	0.996	0.996
1.154	1.154	1.154	1.154	1.154	1.053	1.053	1.053	1.053	1.053	0.996	0.996	0.996	0.996	0.996
1.154	1.154	1.154	1.154	1.154	1.053	1.053	1.053	1.053	1.053	0.996	0.996	0.996	0.996	0.996
1.094	1.094	1.094	1.094	1.094	0.994	0.994	0.994	0.994	0.994	0.937	0.937	0.937	0.937	0.937
1.094	1.094	1.094	1.094	1.094	0.994	0.994	0.994	0.994	0.994	0.937	0.937	0.937	0.937	0.937
1.094	1.094	1.094	1.094	1.094	0.994	0.994	0.994	0.994	0.994	0.937	0.937	0.937	0.937	0.937
1.094	1.094	1.094	1.094	1.094	0.994	0.994	0.994	0.994	0.994	0.937	0.937	0.937	0.937	0.937
1.094	1.094	1.094	1.094	1.094	0.994	0.994	0.994	0.994	0.994	0.937	0.937	0.937	0.937	0.937
1.094	1.094	1.094	1.094	1.094	0.994	0.994	0.994	0.994	0.994	0.937	0.937	0.937	0.937	0.937
1.094	1.094	1.094	1.094	1.094	0.994	0.994	0.994	0.994	0.994	0.937	0.937	0.937	0.937	0.937
1.094	1.094	1.094	1.094	1.094	0.994	0.994	0.994	0.994	0.994	0.937	0.937	0.937	0.937	0.937

GUIDEWIRE Page 11

AN UNREALISTIC ILLUSTRATION



ENSEMBLE Models 1 - 5

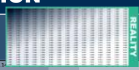
Sum of the squared error = 8.47

1.278	1.278	1.278	1.278	1.259	1.218	1.218	1.164	1.164	1.164	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
1.278	1.278	1.278	1.278	1.259	1.218	1.218	1.164	1.164	1.164	1.141	1.141	1.063	1.063	1.063
1.278	1.278	1.278	1.278	1.259	1.218	1.218	1.164	1.164	1.164	1.141	1.141	1.063	1.063	1.063
1.278	1.278	1.278	1.278	1.259	1.218	1.218	1.164	1.164	1.164	1.141	1.141	1.063	1.063	1.063
1.278	1.278	1.278	1.278	1.259	1.218	1.218	1.164	1.164	1.164	1.141	1.141	1.063	1.063	1.063
1.278	1.278	1.278	1.278	1.259	1.218	1.218	1.164	1.164	1.164	1.141	1.141	1.063	1.063	1.063
1.278	1.278	1.278	1.278	1.259	1.218	1.218	1.164	1.164	1.164	1.141	1.141	1.063	1.063	1.063
1.278	1.278	1.278	1.278	1.259	1.218	1.218	1.164	1.164	1.164	1.141	1.141	1.063	1.063	1.063
1.230	1.230	1.230	1.230	1.211	1.170	1.170	1.116	1.116	1.116	1.093	1.093	1.015	1.015	1.015
1.230	1.230	1.230	1.230	1.211	1.170	1.170	1.116	1.116	1.116	1.093	1.093	1.015	1.015	1.015
1.172	1.172	1.172	1.172	1.152	1.058	1.058	1.058	1.058	1.058	1.015	1.015	1.015	1.015	0.975
1.172	1.172	1.172	1.172	1.112	1.018	1.018	1.018	1.018	1.018	0.975	0.975	0.975	0.975	0.975
1.172	1.172	1.172	1.172	1.112	1.018	1.018	1.018	1.018	1.018	0.975	0.975	0.975	0.975	0.975
1.172	1.172	1.172	1.172	1.112	1.018	1.018	1.018	1.018	1.018	0.975	0.975	0.975	0.975	0.975
1.172	1.172	1.172	1.172	1.172	1.054	0.994	0.994	0.974	0.974	0.952	0.952	0.952	0.952	0.952
1.172	1.172	1.172	1.172	1.172	1.054	0.994	0.994	0.974	0.974	0.952	0.952	0.952	0.952	0.952
1.148	1.148	1.148	1.148	1.070	1.030	0.970	0.970	0.951	0.951	0.928	0.928	0.928	0.928	0.928
1.148	1.148	1.148	1.148	1.070	1.030	0.970	0.970	0.951	0.951	0.928	0.928	0.928	0.928	0.928
1.148	1.148	1.148	1.148	1.070	1.030	0.970	0.970	0.951	0.951	0.928	0.928	0.928	0.928	0.928
1.148	1.148	1.148	1.148	1.070	1.030	0.970	0.970	0.951	0.951	0.928	0.928	0.928	0.928	0.928
1.148	1.148	1.148	1.148	1.070	1.030	1.030	0.970	0.951	0.951	0.928	0.928	0.928	0.928	0.928
1.070	1.070	1.070	1.070	1.070	1.030	1.030	0.970	0.951	0.951	0.928	0.928	0.928	0.928	0.928
1.070	1.070	1.070	1.070	1.070	1.030	1.030	0.970	0.951	0.951	0.928	0.928	0.928	0.928	0.928
1.070	1.070	1.070	1.070	1.070	1.030	1.030	0.970	0.951	0.951	0.928	0.928	0.928	0.928	0.928

GUIDEWIRE Page 12

AN UNREALISTIC ILLUSTRATION

ENSEMBLE
Models 1 - 9



1.296	1.296	1.296	1.296	1.285	1.262	1.262	1.191	1.191	1.152	1.14	[illegible]	[illegible]	[illegible]	[illegible]
1.296	1.296	1.296	1.296	1.285	1.262	1.262	1.232	1.191	1.152	1.140	1.140	1.096	1.053	1.053
1.296	1.296	1.296	1.296	1.285	1.262	1.262	1.232	1.232	1.152	1.140	1.140	1.096	1.053	1.053
1.296	1.296	1.296	1.296	1.285	1.262	1.262	1.232	1.232	1.193	1.140	1.140	1.096	1.053	1.053
1.296	1.296	1.296	1.296	1.285	1.262	1.262	1.232	1.196	1.148	1.136	1.095	1.009	1.009	1.009
1.296	1.296	1.296	1.296	1.285	1.262	1.262	1.232	1.196	1.148	1.136	1.095	1.009	1.009	1.009
1.296	1.296	1.296	1.296	1.285	1.262	1.262	1.232	1.196	1.148	1.136	1.101	1.009	1.009	1.009
1.296	1.296	1.296	1.296	1.285	1.262	1.262	1.232	1.196	1.148	1.136	1.101	1.009	1.009	1.009
1.269	1.269	1.269	1.269	1.258	1.236	1.236	1.169	1.169	1.122	1.076	1.042	0.992	0.992	0.982
1.269	1.269	1.269	1.269	1.258	1.236	1.236	1.169	1.130	1.122	1.076	1.042	0.992	0.992	0.982
1.237	1.237	1.237	1.237	1.226	1.140	1.140	1.103	1.064	1.056	0.998	0.998	0.992	0.992	0.960
1.237	1.237	1.237	1.237	1.203	1.118	1.118	1.042	1.042	1.034	0.976	0.976	0.970	0.970	0.960
1.166	1.166	1.166	1.166	1.133	1.047	1.008	1.008	1.008	1.000	0.976	0.976	0.970	0.970	0.960
1.124	1.166	1.166	1.166	1.133	1.047	1.005	1.005	1.008	1.000	0.966	0.966	0.960	0.960	0.960
1.091	1.091	1.133	1.133	1.133	1.026	0.992	0.992	0.984	0.976	0.953	0.947	0.947	0.947	0.947
1.091	1.091	1.091	1.133	1.091	1.026	0.992	0.992	0.981	0.976	0.953	0.947	0.947	0.947	0.947
1.078	1.078	1.078	1.078	1.035	1.012	0.979	0.979	0.968	0.960	0.938	0.931	0.931	0.933	0.933
1.078	1.078	1.078	1.078	1.035	1.012	0.979	0.979	0.968	0.960	0.931	0.931	0.931	0.933	0.933
1.078	1.078	1.078	1.078	1.035	1.012	0.979	0.979	0.968	0.960	0.931	0.931	0.931	0.933	0.933
1.078	1.078	1.078	1.078	1.035	1.012	0.979	0.979	0.968	0.954	0.931	0.931	0.931	0.933	0.933
1.078	1.078	1.078	1.078	1.035	1.003	1.003	0.969	0.952	0.944	0.931	0.931	0.931	0.933	0.933
1.035	1.035	1.035	1.035	1.035	1.003	1.003	0.963	0.952	0.944	0.931	0.931	0.931	0.933	0.933
1.035	1.035	1.035	1.035	1.035	1.003	0.996	0.963	0.952	0.944	0.931	0.931	0.931	0.933	0.933
1.035	1.035	1.035	1.035	1.035	0.996	0.996	0.963	0.952	0.944	0.931	0.931	0.931	0.933	0.933

Sum of the squared error = 7.35

GUIDEWIRE

 Page 13

Ensembles

"Ensemble modeling has taken the [Predictive Analytics] industry by storm.

It's often considered the most important predictive modeling advancement of this century's first decade."

Siegel, E. (2013). *Predictive Analytics.*

GUIDEWIRE

 Page 14

Basic Approaches – Bagging and Boosting

Basics of Ensembles

How do you take one set of data and one modeling method and get multiple models?!

1. **Data**
2. **Modeling technique(s)**
3. **Method for combining models**

Basics of Ensembles

Remember our credibility-weighting of statewide and countrywide averages?

1. **We get variety from using different data.**
2. **Only one technique is used (averaging).**
3. **We combine through n/(n+k).**

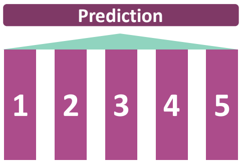
Basics of Ensembles

Bagging = Bootstrap aggregation

- One modeling technique is used on several randomly sampled versions of the data.
- Bootstrapped datasets are built by sampling with replacement to build several equal size datasets.

Component models within an ensemble are "learners."

Basics of Ensembles



Individual learners stand side-by-side.
Weighting can be applied to the average.

Bagging

With learners built on different versions of the data, bagging averages predicted estimates together, thereby reducing the variance of the prediction.

Basics of Ensembles

Adaboost (short for adaptive boosting) is one of the original versions of boosting.

Predictions from the first learner are compared to actuals. Misclassified instances are given more weight ("boosted") in subsequent learners. Later learners have a chance to explicitly correct errors from previous ones.

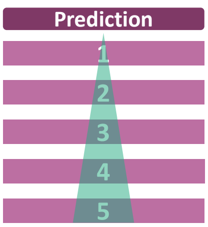
Letting subsequent models focus on the residuals of prior models is the essence of a boosting approach.

Basics of Ensembles

Boosting

- Approach to the data is modified, not the data itself.
- Boosting is effective at reducing the bias of the prediction.

Learners layer on top of each other.
Subsequent learners take into account the results of prior learners.



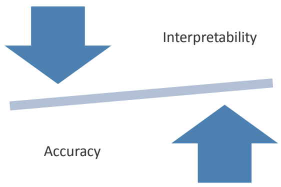
Complexity – Issues and Advantages

Accuracy and Interpretability

We often frame our thoughts as a trade-off between a better prediction versus how well we can explain it.

"…the product team needs to weigh the benefit of the added lift compared to the need for transparency."

- Jan/Feb 2017 issue of Actuarial Review, p.31



Accuracy and Interpretability

"Framing the question as the choice between accuracy and interpretability is an incorrect interpretation of what the goal of a statistical analysis is.

The point of a model is to get useful information about the relation between the response and predictor variables. Interpretability is a way of getting information."

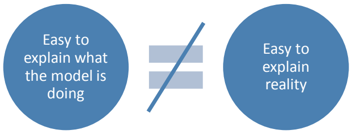
Breiman, L. (2001). Statistical Modeling: The Two Cultures. *Statistical Science*, Vol. 16, No. 3.

Complexity and Interpretability

Conventional Wisdom → GLMs and single trees are easy to explain. Machine learning techniques are not.

"…greater sophistication also makes the reasons behind the results less transparent and harder to explain."

- Jan/Feb 2017 issue of Actuarial Review, p.32



Breiman again…

"…when a model is fit to data to draw quantitative conclusions…the conclusions are about the model's mechanism, not about nature's mechanism.

It follows that…if the model is a poor emulation of nature, the conclusions may be wrong.

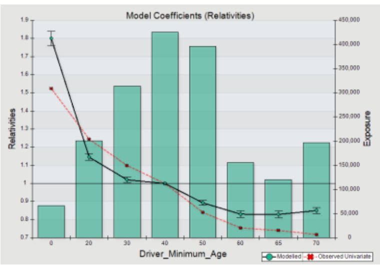
These truisms have often been ignored…It is a strange phenomenon – once a model is made, then it becomes truth and the conclusions from it are infallible."

Breiman, L. (2001). Statistical Modeling: The Two Cultures. *Statistical Science*, Vol. 16, No. 3.

Complexity and Interpretability

It is easy to explain that the GLM identifies youthful drivers as having higher frequency.

But how do we quantify the full frequency of the group of youthful drivers?



Complexity and Interpretability

GLM relativities are useful for identifying rating factors to be used in conjunction with other rating factors. They are harder to interpret as fundamental truths about risk levels.

Reality doesn't have youthful drivers without correlations with other fields – territories, credit no-hit, etc.

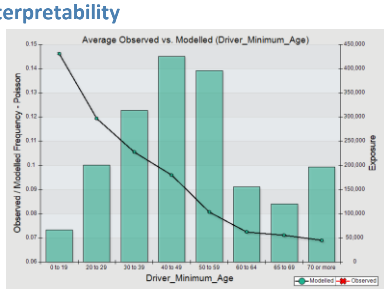
Even slight aliasing can distort the relativities. GLMs are somewhat arbitrary in how they assigned signal to its different predictors. They do it in an internally consistent way to optimize the fitting function, but as was noted earlier, different allocations can be almost equally as valid.

Complexity and Interpretability

We often look at Observed versus Modeled charts.

This checks the balance of the GLM model.

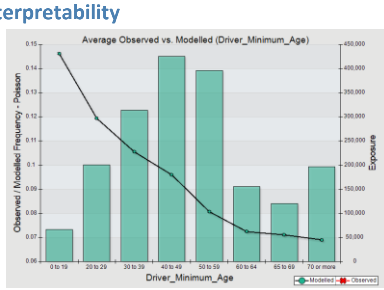
It also shows that, all things considered, youthful is a higher group.



Complexity and Interpretability

Charts showing observed values against modeled predictions do not depend on the model being a GLM.

OvM charts are good for checking *and explaining* any model – ensemble of trees, neural nets, SVM, GLM, etc.



Complexity and Reality

Reality exhibits both broad trends (youthfuls are higher frequency) and complex relationships.

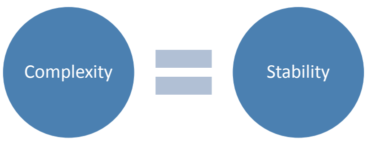
"Complex models" put complex interactions into their inner workings because it fits reality better.

Using a complex model does not change the broad trends – they can still be identified and represented.

Complexity and Stability

Simple models depend on fewer fields. If those fields change…

Complex models exhibit less dependence on the exact value of individual fields.



Complexity and Stability

The table to the right shows the impact on the model output (999-point range) given increases in the most important field.

This is from a Workers Compensation example; the changing field was payroll.

Score Difference	Volume %
-150	0%
-140	0%
-130	0%
-120	0%
-110	0%
-100	1%
-90	3%
-80	1%
-70	3%
-60	1%
-50	2%
-40	2%
-30	1%
-20	3%
-10	1%
0	56%
10	6%
20	4%
30	3%
40	2%
50	4%
60	1%
70	1%
80	0%
90	1%
100	1%
110	1%
120	0%

Combining Linear Regression and Ensembles

Case Study

Worker's Compensation data from 2004 thru 2016Q2
Exposures represent $100,000 in payroll

Frequency target

Training Data:	70% of 2004-2013 data, selected at random
Validation Data:	30% of 2004-2013 data, the balance of this group
Test Data:	2014 and 2015 data

All results here are shown on the Test data

GUIDEWIRE

Case Study

Two modeling methodologies are used.

- A forward stepwise GLM targeting a collection of 30 possible predictors.
- A boosted ensemble of trees using the same collection of 30 possible predictors. Analogous to the forward stepwise GLM, an automated process was used to select the primary model parameters of learning rate and tree depth.

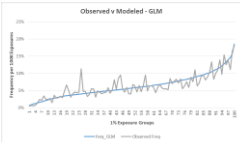
In both cases, modeler discretion was limited to the number of iterations. The assumption here is that both techniques could be improved by human intervention.

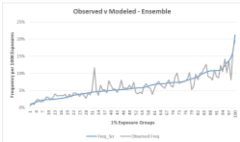
GUIDEWIRE

Case Study – GLM versus Ensemble

How do these methods compare when simply building a "ground-up" frequency model? On the surface, similar lift and fit.

	GLM	Ensemble
Min	0.7%	0.9%
Max	18.5%	21.2%
Lift	26.3	22.5
Spread	0.178	0.203





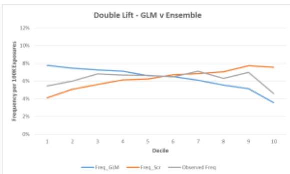
GUIDEWIRE

Case Study – GLM versus Ensemble

A double lift chart shows mixed results as well.

However, is this comparison valid?

Is this the proper way to take advantage of the particular strengths and weaknesses of each approach?



GUIDEWIRE

Combining Linear Regression and Ensembles

We often think about the linear and non-linear signal in the data.

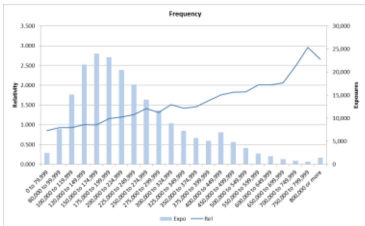
	(log) Linear	Non-linear, Combinatorial
GLM	Efficient representation	Possible (to a degree) to represent, but cumbersome to explore
Ensembles of Trees	Inefficient representation	Natural representation and exploration

GUIDEWIRE

Combining Linear Regression and Ensembles

When there is linear signal, a GLM represents this in a straight-forward manner.

Imagine what it would take for a tree to represent this same information…

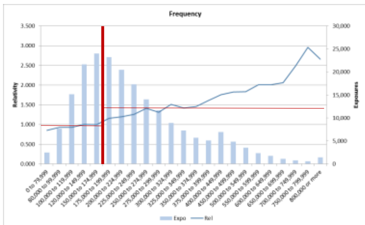


GUIDEWIRE

Combining Linear Regression and Ensembles

When there is linear signal, a GLM represents this in a straight-forward manner.

Imagine what it would take for a tree to represent this same information…

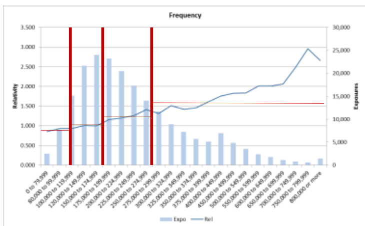


GUIDEWIRE

Combining Linear Regression and Ensembles

A tree-based approach would have to go several layers deep to even approximate the information in the GLM for this linear relationship.

This is inefficient.

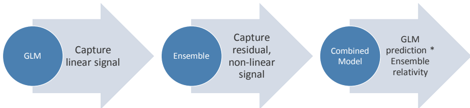


GUIDEWIRE

Combining Linear Regression and Ensembles

This isn't a competition. We should combine methods in ways that enhance their strengths and limit their weaknesses.

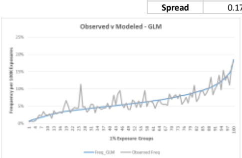
The first approach we'll try is to build a GLM and then model the residuals using the Ensemble.

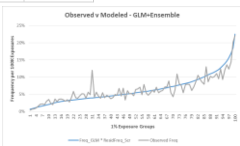


Case Study – GLM versus GLM+Ensemble

The predictions from the Ensemble add noticeable and consistent lift to the model. Ensemble relativities ranged from +64% to -39%.

	GLM	GLM+Ensemble
Min	0.7%	0.7%
Max	18.5%	22.5%
Lift	26.3	33.3
Spread	0.178	0.218



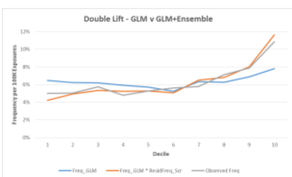


Case Study – GLM versus GLM+Ensemble

A double lift chart shows a clearly better result as well.

Specifically in the cases where the combined model and the GLM disagree, the combined models is consistently and dramatically more accurate.

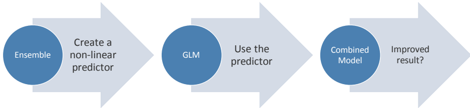
Remember that these results are on a pure Test dataset.



Combining Linear Regression and Ensembles

What if we let the Ensemble go first instead?

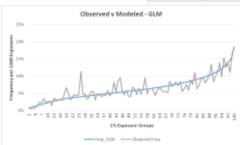
Part of the Ensemble output for the approach we used presents the model prediction as a 3-digit score. This Score was attached to the data and considered as an additional predictor representing the non-linear signal in the data.

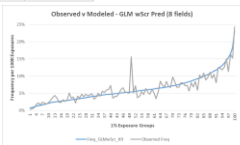


Case Study – GLM versus GLM with non-linear predictor

Like the other combined approach, the lift of the model is noticeably improved.

	GLM	GLM wScr Pred
Min	0.7%	0.8%
Max	18.5%	23.1%
Lift	26.3	30.8
Spread	0.178	0.224

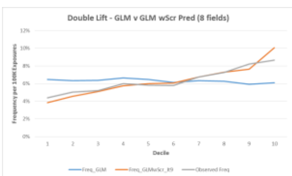




Case Study – GLM versus GLM with non-linear predictor

And again, a double lift chart shows a clearly better result as well.

Specifically in the cases where the combined model and the GLM disagree, the combined models is consistently and dramatically more accurate.



Case Study – GLM versus GLM with non-linear predictor

It is interesting to examine the output of the forward stepwise procedure for the base GLM and the GLM with the non-linear predictor.

Baseline GLM	
Variable(s) Added	Deviance
NULL MODEL	18,402
GoverningClassCode	17,830
EEA_Policy_ClaimCount_Prior3Years	17,548
EEA_Policy_TotalPayroll	17,148
EEA_Policy_ChangeInClassCodesState_Ind	17,019
EEA_Policy_OfficeAndClericalManualPremium_Pct	16,763
EEA_Policy_SubcontractorsManualPremium_Pct	16,670
CoverageState	16,640
EEA_NAICS_First2Digits	16,584

GLM with non-linear predictor	
Variable(s) Added	Deviance
NULL MODEL	18,402
Scr_Freq_f6bdf	16,648
GoverningClassCode	16,486
EEA_Policy_ClaimCount_PriorYear	16,466
CoverageState	16,439
EEA_Policy_TotalPayroll	16,407
EEA_Policy_OfficeAndClericalManualPremium_Pct	16,373
EEA_Policy_ClaimCount_Prior3Years	16,370
Deductible	16,357

Case Study – Combined versus Combined

Is there a performance difference in the two combined model approaches? Not on the basis of lift.

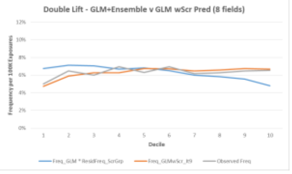
It is notable that the creation of a non-linear predictor serves to simplify the entire model. The same lift is achieved with the loss of fewer degrees of freedom.

	GLM+Ensemble	GLM wScr Pred
Min	0.7%	0.8%
Max	22.5%	23.1%
Lift	33.3	30.8
Spread	0.218	0.224
# Levels	76	70
dF	67	62
Price Points	27,417,600	5,140,800

Case Study – Combined versus Combined

The double lift chart in this case shows a clear winner.

Despite being a simpler model, when the two approaches disagree the GLM which uses a non-linear predictor is consistently more accurate than a GLM plus a refinement based on a residual Ensemble model.



Case Study – Combined versus Combined

Is there really a clear winner?

In the case of Pricing, there are distinct advantages to modeling the residuals of a baseline GLM.

- By taking the GLM results as a given, the "complicated" model produces a single rate adjustment factor.
- The combined model still looks like a traditional rating plan.
- The Ensemble-based adjustment factor can be considered on its own terms – acceptability to agents, customers, regulators, etc.

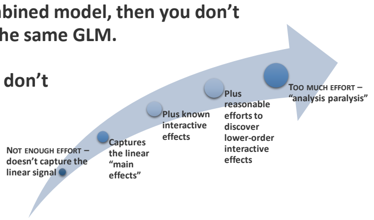
Also, we should note this is one result for one target on one dataset for one line of business.

GUIDEWIRE

GLM within a combined approach

It is important to note that if you know from the beginning you are building a combined model, then you don't necessarily build the same GLM.

Combined models don't necessarily take more time.



GUIDEWIRE

Summary

- **Ensembles work by combining information from multiple models.**
- **Bagging averages predictions; boosting focuses on residuals.**
- **GLMs parse effects to individual fields. The question of who has a high or low prediction is different.**
- **Observed versus Modeled graphs are independent of modeling method. They can be used to explain complex models.**
- **Reality, with its simple trends *and* complexity exists without regard to our modeling method.**
- **There is great potential to combine modeling methods.**

GUIDEWIRE

Questions?

Christopher Cooksey, FCAS, MAAA
Head Actuary, Data and Analytics

Guidewire Software

ccooksey@guidewire.com