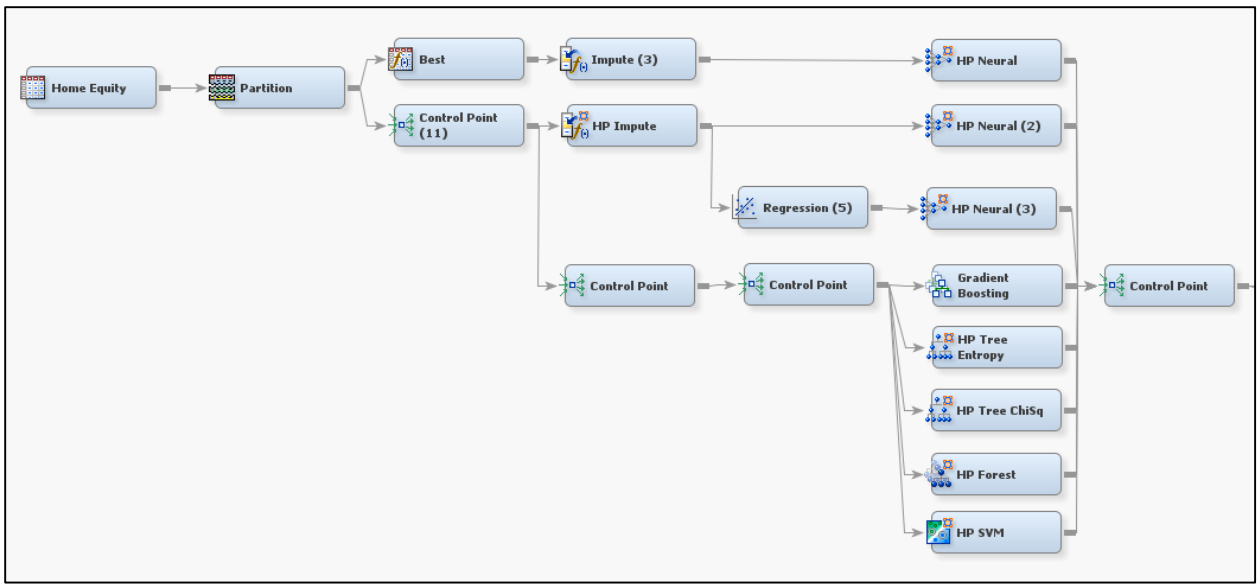
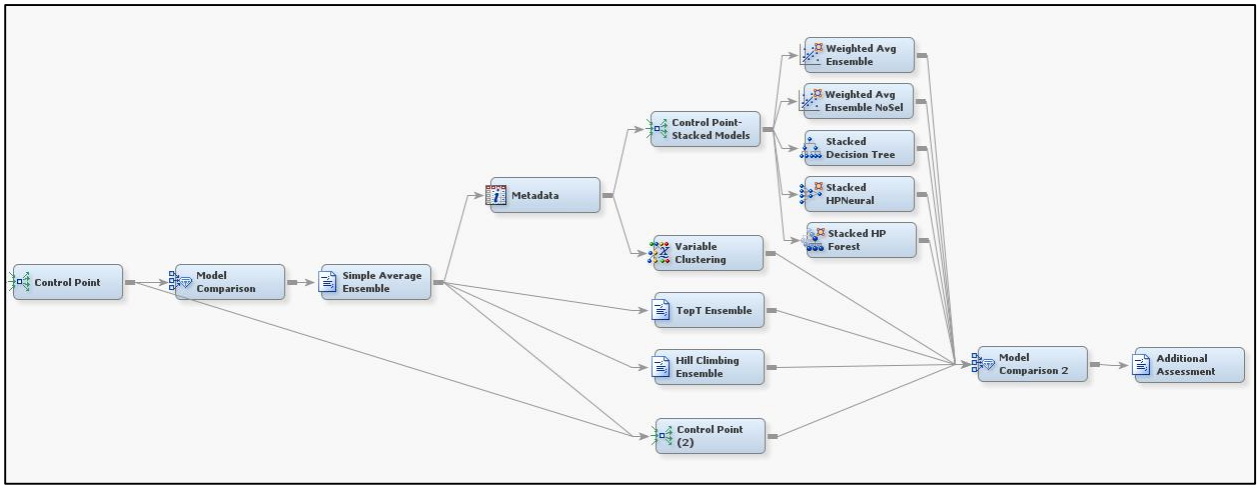
Ensemble Modeling in SAS® Enterprise Miner™

Please see the SAS Global Forum paper Ensemble Modeling: Recent Advances and Applications for details on the SAS Enterprise Miner flows provided here that implement the following ensemble methods that take model performance into account: top-*t*, hill-climbing, clustering-based selection, and stacking methods. Two XML files are provided:



EnsembleFullflow.xml contains an entire predictive modeling and ensemble flow, comprising the "Common Practices" flow from the paper (shown above) connected to the ensemble subflow (shown below) so you can see, and run, the whole process.



EnsembleSubflow.xml has just the ensemble portion of the flow that you can connect to an existing predictive modeling flow. After importing this XML file into your project, you can copy the entire flow into the diagram that has your predictive modeling flow, connect the flows together, and run.

See the README file for instructions on how to import these XML files and quickly get started with these more sophisticated ensemble methods.

Note there are several nodes that directly create ensemble models in SAS Enterprise Miner that have been covered in previous SAS Global Forum papers:

- The **Ensemble** node for simple averaging/voting/maximum of multiple models
- The **Start Group** and **End Group** nodes for bagging and boosting
- The **Gradient Boosting** node and **HP Forest** node for tree-based methods

See Leveraging Ensemble Models in SAS Enterprise Miner and The Power of the Group Processing Facility in SAS Enterprise Miner for more information.