<?xml version="1.0" encoding="UTF-8"?><process version="10.3.001">

<context>

<input/>

<output/>

<macros/>

</context>

<operator activated="true" class="process" compatibility="9.4.000" expanded="true" name="Process" origin="GENERATED\_SAMPLE">

<parameter key="logverbosity" value="init"/>

<parameter key="random\_seed" value="2001"/>

<parameter key="send\_mail" value="never"/>

<parameter key="notification\_email" value=""/>

<parameter key="process\_duration\_for\_mail" value="30"/>

<parameter key="encoding" value="SYSTEM"/>

<process expanded="true">

<operator activated="true" class="retrieve" compatibility="10.3.001" expanded="true" height="68" name="Retrieve Expanded\_data\_with\_more\_features" width="90" x="45" y="85">

<parameter key="repository\_entry" value="//Project\_1/Expanded\_data\_with\_more\_features"/>

</operator>

<operator activated="true" class="subprocess" compatibility="10.3.001" expanded="true" height="124" name="Preprocessing &amp; EDA" width="90" x="179" y="34">

<process expanded="true">

<operator activated="false" class="replace\_missing\_values" compatibility="10.3.001" expanded="true" height="103" name="Replace Missing Values" width="90" x="313" y="238">

<parameter key="return\_preprocessing\_model" value="false"/>

<parameter key="attribute\_filter\_type" value="all"/>

<parameter key="attribute" value=""/>

<parameter key="attributes" value=""/>

<parameter key="use\_except\_expression" value="false"/>

<parameter key="value\_type" value="attribute\_value"/>

<parameter key="use\_value\_type\_exception" value="false"/>

<parameter key="except\_value\_type" value="time"/>

<parameter key="block\_type" value="attribute\_block"/>

<parameter key="use\_block\_type\_exception" value="false"/>

<parameter key="except\_block\_type" value="value\_matrix\_row\_start"/>

<parameter key="invert\_selection" value="false"/>

<parameter key="include\_special\_attributes" value="false"/>

<parameter key="default" value="average"/>

<list key="columns"/>

</operator>

<operator activated="true" class="numerical\_to\_polynominal" compatibility="10.3.001" expanded="true" height="82" name="Numerical to Polynominal" width="90" x="45" y="34">

<parameter key="attribute\_filter\_type" value="single"/>

<parameter key="attribute" value="NrSiblings"/>

<parameter key="attributes" value=""/>

<parameter key="use\_except\_expression" value="false"/>

<parameter key="value\_type" value="numeric"/>

<parameter key="use\_value\_type\_exception" value="false"/>

<parameter key="except\_value\_type" value="real"/>

<parameter key="block\_type" value="value\_series"/>

<parameter key="use\_block\_type\_exception" value="false"/>

<parameter key="except\_block\_type" value="value\_series\_end"/>

<parameter key="invert\_selection" value="false"/>

<parameter key="include\_special\_attributes" value="false"/>

</operator>

<operator activated="true" class="replace" compatibility="10.3.001" expanded="true" height="82" name="Replace" width="90" x="45" y="136">

<parameter key="attribute\_filter\_type" value="single"/>

<parameter key="attribute" value="WklyStudyHours"/>

<parameter key="attributes" value=""/>

<parameter key="use\_except\_expression" value="false"/>

<parameter key="value\_type" value="nominal"/>

<parameter key="use\_value\_type\_exception" value="false"/>

<parameter key="except\_value\_type" value="file\_path"/>

<parameter key="block\_type" value="single\_value"/>

<parameter key="use\_block\_type\_exception" value="false"/>

<parameter key="except\_block\_type" value="single\_value"/>

<parameter key="invert\_selection" value="false"/>

<parameter key="include\_special\_attributes" value="false"/>

<parameter key="replace\_what" value="10-May"/>

<parameter key="replace\_by" value="5-10"/>

</operator>

<operator activated="true" class="blending:set\_role" compatibility="10.3.001" expanded="true" height="82" name="Set Role" width="90" x="179" y="136">

<list key="set\_roles">

<parameter key="att1" value="id"/>

<parameter key="MathScore" value="label"/>

</list>

</operator>

<operator activated="true" class="filter\_examples" compatibility="10.3.001" expanded="true" height="103" name="Filter Examples" width="90" x="313" y="85">

<parameter key="parameter\_expression" value=""/>

<parameter key="condition\_class" value="no\_missing\_attributes"/>

<parameter key="invert\_filter" value="false"/>

<list key="filters\_list"/>

<parameter key="filters\_logic\_and" value="true"/>

<parameter key="filters\_check\_metadata" value="true"/>

</operator>

<operator activated="true" class="model\_simulator:data\_statistics" compatibility="10.3.001" expanded="true" height="82" name="Statistics" width="90" x="447" y="85"/>

<operator activated="true" class="concurrency:correlation\_matrix" compatibility="10.3.001" expanded="true" height="103" name="Correlation Matrix" width="90" x="447" y="238">

<parameter key="type" value="include attributes"/>

<parameter key="attribute\_filter\_type\_" value="all attributes"/>

<parameter key="select\_attribute" value=""/>

<parameter key="select\_subset" value=""/>

<parameter key="also\_apply\_to\_special\_attributes\_(id,\_label..)" value="true"/>

<parameter key="attribute\_filter\_type" value="all"/>

<parameter key="attribute" value=""/>

<parameter key="attributes" value=""/>

<parameter key="use\_except\_expression" value="false"/>

<parameter key="value\_type" value="attribute\_value"/>

<parameter key="use\_value\_type\_exception" value="false"/>

<parameter key="except\_value\_type" value="time"/>

<parameter key="block\_type" value="attribute\_block"/>

<parameter key="use\_block\_type\_exception" value="false"/>

<parameter key="except\_block\_type" value="value\_matrix\_row\_start"/>

<parameter key="invert\_selection" value="false"/>

<parameter key="include\_special\_attributes" value="false"/>

<parameter key="normalize\_weights" value="true"/>

<parameter key="squared\_correlation" value="false"/>

</operator>

<operator activated="true" class="blending:select\_attributes" compatibility="10.3.001" expanded="true" height="82" name="Select Attributes" width="90" x="581" y="187">

<parameter key="type" value="exclude attributes"/>

<parameter key="attribute\_filter\_type" value="a subset"/>

<parameter key="select\_attribute" value="WritingScore"/>

<parameter key="select\_subset" value="att1␞EthnicGroup␞Gender␞IsFirstChild␞LunchType␞NrSiblings␞ParentEduc␞ParentMaritalStatus␞PracticeSport␞TestPrep␞TransportMeans␞WklyStudyHours␞WritingScore"/>

<parameter key="also\_apply\_to\_special\_attributes\_(id,\_label..)" value="false"/>

</operator>

<operator activated="true" class="nominal\_to\_numerical" compatibility="10.3.001" expanded="true" height="103" name="Nominal to Numerical" width="90" x="715" y="187">

<parameter key="return\_preprocessing\_model" value="false"/>

<parameter key="attribute\_filter\_type" value="all"/>

<parameter key="attribute" value=""/>

<parameter key="attributes" value=""/>

<parameter key="use\_except\_expression" value="false"/>

<parameter key="value\_type" value="nominal"/>

<parameter key="use\_value\_type\_exception" value="false"/>

<parameter key="except\_value\_type" value="file\_path"/>

<parameter key="block\_type" value="single\_value"/>

<parameter key="use\_block\_type\_exception" value="false"/>

<parameter key="except\_block\_type" value="single\_value"/>

<parameter key="invert\_selection" value="false"/>

<parameter key="include\_special\_attributes" value="false"/>

<parameter key="coding\_type" value="dummy coding"/>

<parameter key="use\_comparison\_groups" value="false"/>

<list key="comparison\_groups"/>

<parameter key="unexpected\_value\_handling" value="all 0 and warning"/>

<parameter key="use\_underscore\_in\_name" value="false"/>

</operator>

<operator activated="true" class="normalize" compatibility="10.3.001" expanded="true" height="103" name="Normalize" width="90" x="849" y="187">

<parameter key="return\_preprocessing\_model" value="false"/>

<parameter key="attribute\_filter\_type" value="all"/>

<parameter key="attribute" value=""/>

<parameter key="attributes" value=""/>

<parameter key="use\_except\_expression" value="false"/>

<parameter key="value\_type" value="numeric"/>

<parameter key="use\_value\_type\_exception" value="false"/>

<parameter key="except\_value\_type" value="real"/>

<parameter key="block\_type" value="value\_series"/>

<parameter key="use\_block\_type\_exception" value="false"/>

<parameter key="except\_block\_type" value="value\_series\_end"/>

<parameter key="invert\_selection" value="false"/>

<parameter key="include\_special\_attributes" value="false"/>

<parameter key="method" value="Z-transformation"/>

<parameter key="min" value="0.0"/>

<parameter key="max" value="1.0"/>

<parameter key="allow\_negative\_values" value="false"/>

</operator>

<connect from\_port="in 1" to\_op="Numerical to Polynominal" to\_port="example set input"/>

<connect from\_op="Numerical to Polynominal" from\_port="example set output" to\_op="Replace" to\_port="example set input"/>

<connect from\_op="Replace" from\_port="example set output" to\_op="Set Role" to\_port="example set input"/>

<connect from\_op="Set Role" from\_port="example set output" to\_op="Filter Examples" to\_port="example set input"/>

<connect from\_op="Filter Examples" from\_port="example set output" to\_op="Statistics" to\_port="example set"/>

<connect from\_op="Statistics" from\_port="statistics" to\_port="out 1"/>

<connect from\_op="Statistics" from\_port="example set" to\_op="Correlation Matrix" to\_port="example set"/>

<connect from\_op="Correlation Matrix" from\_port="example set" to\_op="Select Attributes" to\_port="example set input"/>

<connect from\_op="Correlation Matrix" from\_port="matrix" to\_port="out 2"/>

<connect from\_op="Select Attributes" from\_port="example set output" to\_op="Nominal to Numerical" to\_port="example set input"/>

<connect from\_op="Nominal to Numerical" from\_port="example set output" to\_op="Normalize" to\_port="example set input"/>

<connect from\_op="Normalize" from\_port="example set output" to\_port="out 3"/>

<portSpacing port="source\_in 1" spacing="0"/>

<portSpacing port="source\_in 2" spacing="0"/>

<portSpacing port="sink\_out 1" spacing="0"/>

<portSpacing port="sink\_out 2" spacing="0"/>

<portSpacing port="sink\_out 3" spacing="0"/>

<portSpacing port="sink\_out 4" spacing="0"/>

</process>

</operator>

<operator activated="true" class="shuffle" compatibility="10.3.001" expanded="true" height="82" name="Shuffle" width="90" x="313" y="187">

<parameter key="use\_local\_random\_seed" value="true"/>

<parameter key="local\_random\_seed" value="1992"/>

</operator>

<operator activated="true" class="filter\_example\_range" compatibility="10.3.001" expanded="true" height="82" name="Test set" width="90" x="447" y="238">

<parameter key="first\_example" value="15395"/>

<parameter key="last\_example" value="19243"/>

<parameter key="invert\_filter" value="false"/>

</operator>

<operator activated="true" class="filter\_example\_range" compatibility="10.3.001" expanded="true" height="82" name="Training set" width="90" x="447" y="136">

<parameter key="first\_example" value="1"/>

<parameter key="last\_example" value="15394"/>

<parameter key="invert\_filter" value="false"/>

</operator>

<operator activated="true" class="concurrency:cross\_validation" compatibility="10.3.001" expanded="true" height="145" name="Cross Validation" width="90" x="581" y="85">

<parameter key="split\_on\_batch\_attribute" value="false"/>

<parameter key="leave\_one\_out" value="false"/>

<parameter key="number\_of\_folds" value="10"/>

<parameter key="sampling\_type" value="automatic"/>

<parameter key="use\_local\_random\_seed" value="true"/>

<parameter key="local\_random\_seed" value="1992"/>

<parameter key="enable\_parallel\_execution" value="true"/>

<process expanded="true">

<operator activated="false" class="linear\_regression" compatibility="10.3.001" expanded="true" height="103" name="Linear Regression" width="90" x="112" y="85">

<parameter key="feature\_selection" value="M5 prime"/>

<parameter key="alpha" value="0.05"/>

<parameter key="max\_iterations" value="10"/>

<parameter key="forward\_alpha" value="0.05"/>

<parameter key="backward\_alpha" value="0.05"/>

<parameter key="eliminate\_colinear\_features" value="true"/>

<parameter key="min\_tolerance" value="0.05"/>

<parameter key="use\_bias" value="true"/>

<parameter key="ridge" value="1.0E-8"/>

</operator>

<operator activated="false" class="h2o:generalized\_linear\_model" compatibility="10.3.001" expanded="true" height="124" name="Generalized Linear Model" width="90" x="112" y="187">

<parameter key="family" value="AUTO"/>

<parameter key="link" value="family\_default"/>

<parameter key="solver" value="AUTO"/>

<parameter key="reproducible" value="false"/>

<parameter key="maximum\_number\_of\_threads" value="4"/>

<parameter key="use\_regularization" value="true"/>

<parameter key="lambda\_search" value="false"/>

<parameter key="number\_of\_lambdas" value="0"/>

<parameter key="lambda\_min\_ratio" value="0.0"/>

<parameter key="early\_stopping" value="true"/>

<parameter key="stopping\_rounds" value="3"/>

<parameter key="stopping\_tolerance" value="0.001"/>

<parameter key="standardize" value="true"/>

<parameter key="non-negative\_coefficients" value="false"/>

<parameter key="add\_intercept" value="true"/>

<parameter key="compute\_p-values" value="false"/>

<parameter key="remove\_collinear\_columns" value="false"/>

<parameter key="missing\_values\_handling" value="MeanImputation"/>

<parameter key="max\_iterations" value="0"/>

<parameter key="specify\_beta\_constraints" value="false"/>

<list key="beta\_constraints"/>

<parameter key="max\_runtime\_seconds" value="0"/>

<list key="expert\_parameters"/>

</operator>

<operator activated="true" class="support\_vector\_machine\_linear" compatibility="10.3.001" expanded="true" height="82" name="SVM (Linear)" width="90" x="112" y="340">

<parameter key="kernel\_cache" value="200"/>

<parameter key="C" value="0.0"/>

<parameter key="convergence\_epsilon" value="0.001"/>

<parameter key="max\_iterations" value="100000"/>

<parameter key="scale" value="true"/>

<parameter key="L\_pos" value="1.0"/>

<parameter key="L\_neg" value="1.0"/>

<parameter key="epsilon" value="0.0"/>

<parameter key="epsilon\_plus" value="0.0"/>

<parameter key="epsilon\_minus" value="0.0"/>

<parameter key="balance\_cost" value="false"/>

<parameter key="quadratic\_loss\_pos" value="false"/>

<parameter key="quadratic\_loss\_neg" value="false"/>

</operator>

<connect from\_port="training set" to\_op="SVM (Linear)" to\_port="training set"/>

<connect from\_op="SVM (Linear)" from\_port="model" to\_port="model"/>

<portSpacing port="source\_training set" spacing="0"/>

<portSpacing port="sink\_model" spacing="0"/>

<portSpacing port="sink\_through 1" spacing="0"/>

</process>

<process expanded="true">

<operator activated="true" class="apply\_model" compatibility="10.3.001" expanded="true" height="82" name="Apply Model (1)" width="90" x="45" y="136">

<list key="application\_parameters"/>

</operator>

<operator activated="true" class="performance\_regression" compatibility="10.3.001" expanded="true" height="82" name="Performance (1)" width="90" x="179" y="187">

<parameter key="main\_criterion" value="first"/>

<parameter key="root\_mean\_squared\_error" value="true"/>

<parameter key="absolute\_error" value="false"/>

<parameter key="relative\_error" value="false"/>

<parameter key="relative\_error\_lenient" value="false"/>

<parameter key="relative\_error\_strict" value="false"/>

<parameter key="normalized\_absolute\_error" value="false"/>

<parameter key="root\_relative\_squared\_error" value="false"/>

<parameter key="squared\_error" value="false"/>

<parameter key="correlation" value="false"/>

<parameter key="squared\_correlation" value="false"/>

<parameter key="prediction\_average" value="false"/>

<parameter key="spearman\_rho" value="false"/>

<parameter key="kendall\_tau" value="false"/>

<parameter key="skip\_undefined\_labels" value="true"/>

<parameter key="use\_example\_weights" value="true"/>

</operator>

<connect from\_port="model" to\_op="Apply Model (1)" to\_port="model"/>

<connect from\_port="test set" to\_op="Apply Model (1)" to\_port="unlabelled data"/>

<connect from\_op="Apply Model (1)" from\_port="labelled data" to\_op="Performance (1)" to\_port="labelled data"/>

<connect from\_op="Performance (1)" from\_port="performance" to\_port="performance 1"/>

<portSpacing port="source\_model" spacing="0"/>

<portSpacing port="source\_test set" spacing="0"/>

<portSpacing port="source\_through 1" spacing="0"/>

<portSpacing port="sink\_test set results" spacing="0"/>

<portSpacing port="sink\_performance 1" spacing="0"/>

<portSpacing port="sink\_performance 2" spacing="0"/>

</process>

</operator>

<operator activated="true" class="apply\_model" compatibility="10.3.001" expanded="true" height="82" name="Apply Model (2)" width="90" x="715" y="238">

<list key="application\_parameters"/>

</operator>

<operator activated="true" class="performance\_regression" compatibility="10.3.001" expanded="true" height="82" name="Performance (2)" width="90" x="849" y="340">

<parameter key="main\_criterion" value="first"/>

<parameter key="root\_mean\_squared\_error" value="true"/>

<parameter key="absolute\_error" value="false"/>

<parameter key="relative\_error" value="false"/>

<parameter key="relative\_error\_lenient" value="false"/>

<parameter key="relative\_error\_strict" value="false"/>

<parameter key="normalized\_absolute\_error" value="false"/>

<parameter key="root\_relative\_squared\_error" value="false"/>

<parameter key="squared\_error" value="false"/>

<parameter key="correlation" value="false"/>

<parameter key="squared\_correlation" value="false"/>

<parameter key="prediction\_average" value="false"/>

<parameter key="spearman\_rho" value="false"/>

<parameter key="kendall\_tau" value="false"/>

<parameter key="skip\_undefined\_labels" value="true"/>

<parameter key="use\_example\_weights" value="true"/>

</operator>

<operator activated="true" class="blending:rename" compatibility="10.3.001" expanded="true" height="82" name="Rename" width="90" x="983" y="289">

<list key="rename attributes">

<parameter key="prediction(MathScore)" value="MathScorePredicted"/>

</list>

<parameter key="from\_attribute" value=""/>

<parameter key="to\_attribute" value=""/>

</operator>

<operator activated="true" class="blending:generate\_columns" compatibility="10.3.001" expanded="true" height="82" name="Generate Attributes" width="90" x="1117" y="289">

<list key="function\_descriptions">

<parameter key="residuals" value="(predicted - MathScore)"/>

</list>

<parameter key="keep\_all\_columns" value="true"/>

</operator>

<connect from\_op="Retrieve Expanded\_data\_with\_more\_features" from\_port="output" to\_op="Preprocessing &amp; EDA" to\_port="in 1"/>

<connect from\_op="Preprocessing &amp; EDA" from\_port="out 1" to\_port="result 1"/>

<connect from\_op="Preprocessing &amp; EDA" from\_port="out 3" to\_op="Shuffle" to\_port="example set input"/>

<connect from\_op="Shuffle" from\_port="example set output" to\_op="Training set" to\_port="example set input"/>

<connect from\_op="Shuffle" from\_port="original" to\_op="Test set" to\_port="example set input"/>

<connect from\_op="Test set" from\_port="example set output" to\_op="Apply Model (2)" to\_port="unlabelled data"/>

<connect from\_op="Training set" from\_port="example set output" to\_op="Cross Validation" to\_port="example set"/>

<connect from\_op="Cross Validation" from\_port="model" to\_op="Apply Model (2)" to\_port="model"/>

<connect from\_op="Apply Model (2)" from\_port="labelled data" to\_op="Performance (2)" to\_port="labelled data"/>

<connect from\_op="Apply Model (2)" from\_port="model" to\_port="result 4"/>

<connect from\_op="Performance (2)" from\_port="performance" to\_port="result 2"/>

<connect from\_op="Performance (2)" from\_port="example set" to\_op="Rename" to\_port="example set input"/>

<connect from\_op="Rename" from\_port="example set output" to\_op="Generate Attributes" to\_port="table input"/>

<connect from\_op="Generate Attributes" from\_port="table output" to\_port="result 3"/>

<portSpacing port="source\_input 1" spacing="0"/>

<portSpacing port="sink\_result 1" spacing="0"/>

<portSpacing port="sink\_result 2" spacing="0"/>

<portSpacing port="sink\_result 3" spacing="0"/>

<portSpacing port="sink\_result 4" spacing="0"/>

<portSpacing port="sink\_result 5" spacing="0"/>

</process>

</operator>

</process>