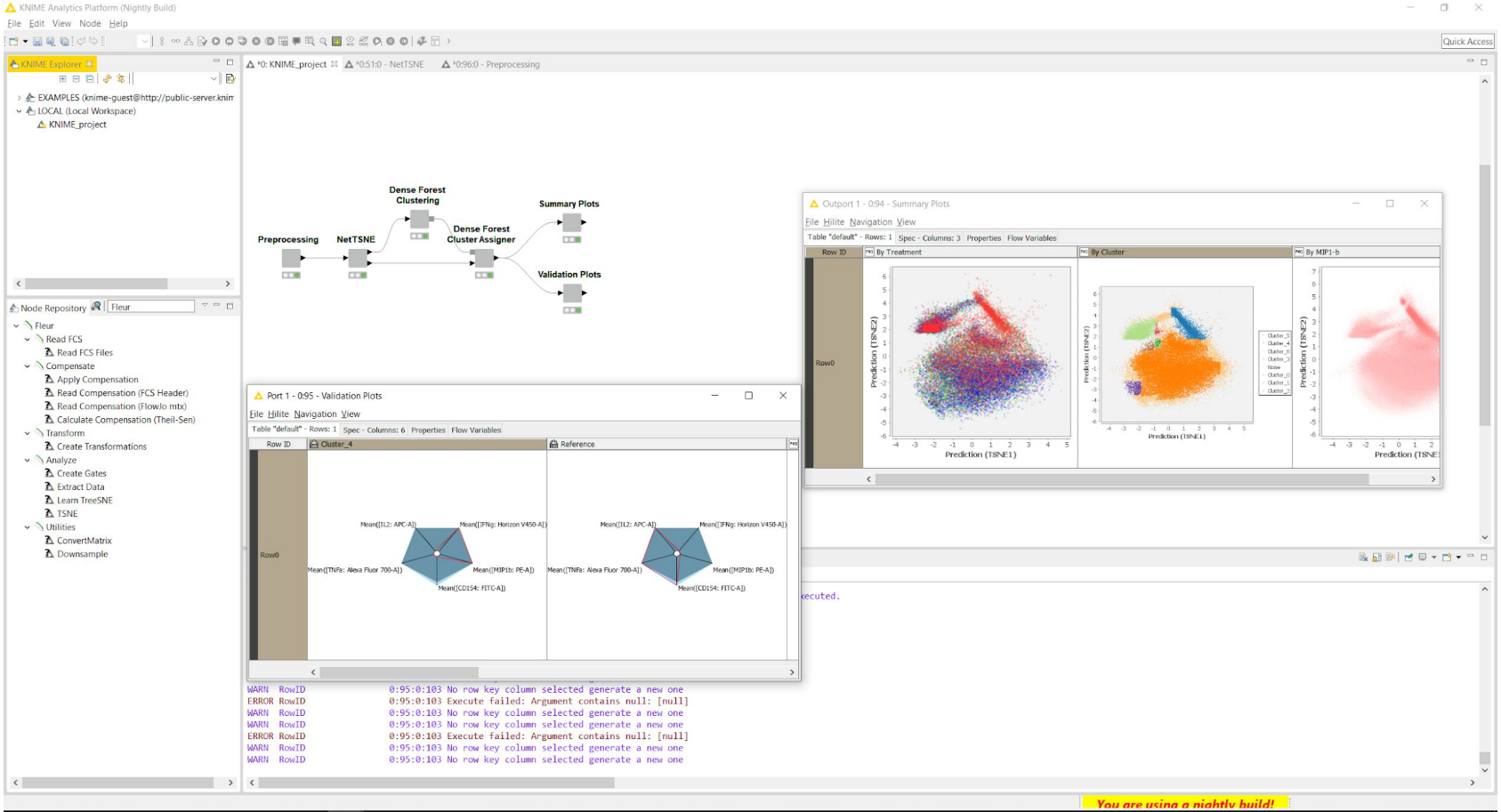
Cytometry Analysis With fleur

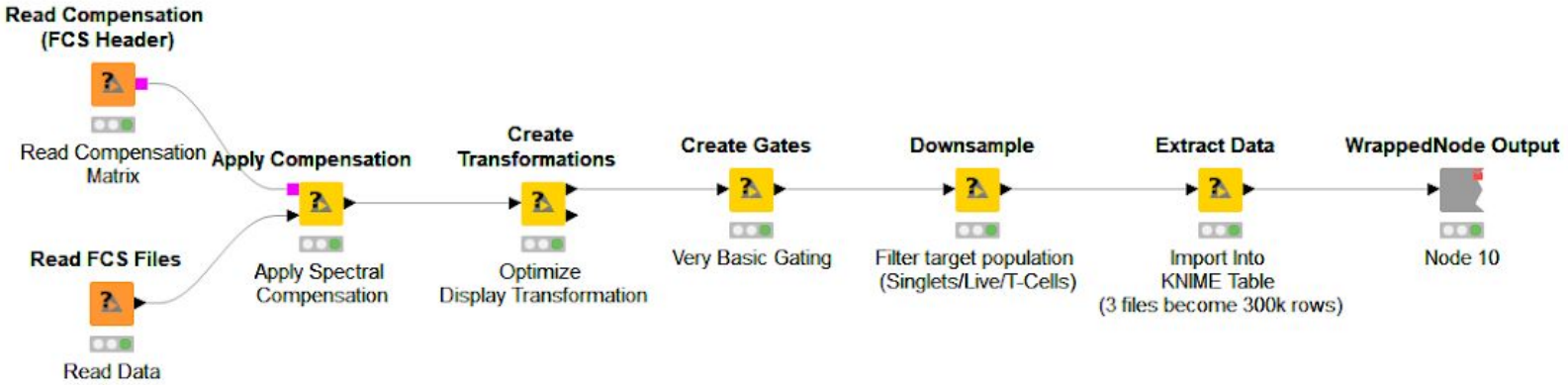
Aaron Hart

Introducing fleur

Why? Modern algorithms are difficult to access with flow cytometry data.



Fleur: KNIME Nodes for Flow Cytometry



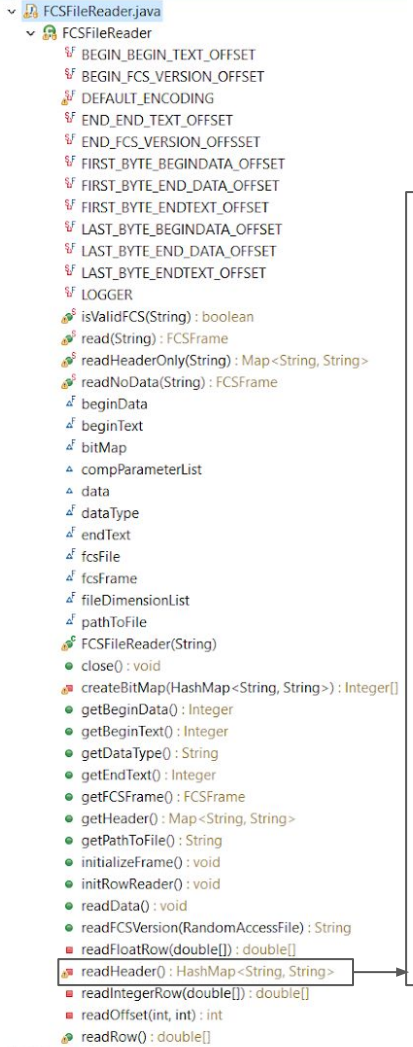
Major Features:

- FCS File Parsing
- Compensation
- Transformation
- Traditional Gating
- Access to wider KNIME

FCS FileReader

Parse a binary file according to FCS3.1 specification.

Read header first for metadata needed to understand the file.




```java
private HashMap<String, String> readHeader() throws IOException {
  // Delimiter is first UTF-8 character in the text section
  final byte[] delimiterBytes = new byte[1];
  fcsFile.seek(beginText);
  fcsFile.read(delimiterBytes);
  final String delimiter = new String(delimiterBytes);

  // Read the rest of the text bytes, this will contain the keywords
  final int textLength = endText - beginText + 1;
  final byte[] keywordBytes = new byte[textLength];

  fcsFile.read(keywordBytes);
  String rawKeywords = new String(keywordBytes, DEFAULT_ENCODING);
  if (rawKeywords.length() > 0
      && rawKeywords.charAt(rawKeywords.length() - 1) == delimiter.charAt(0)) {
    rawKeywords = rawKeywords.substring(0, rawKeywords.length() - 1);
  }
  final StringTokenizer s = new StringTokenizer(rawKeywords, delimiter);
  final HashMap<String, String> header = new HashMap<>();
  Boolean ok = true;
    while (s.hasMoreTokens() && ok) {
      final String key = s.nextToken().trim();
      if (key.trim().isEmpty()) {
        ok = false;
      } else {
        try {
        final String value = s.nextToken().trim();
        header.put(key, value);
        } catch (NoSuchElementException e) {
          String message = "Keyword value for: " + key + " does not exist.  Header appears
          LogFactory.createLogger(this.getClass().getName()).log(Level.FINE, message, e);
        }
      }
    }

  HashFunction md = Hashing.sha256();
  HashCode code = md.hashBytes(keywordBytes);
  header.put("SHA-256", code.toString());
  return header;
}
```

FCSFrame

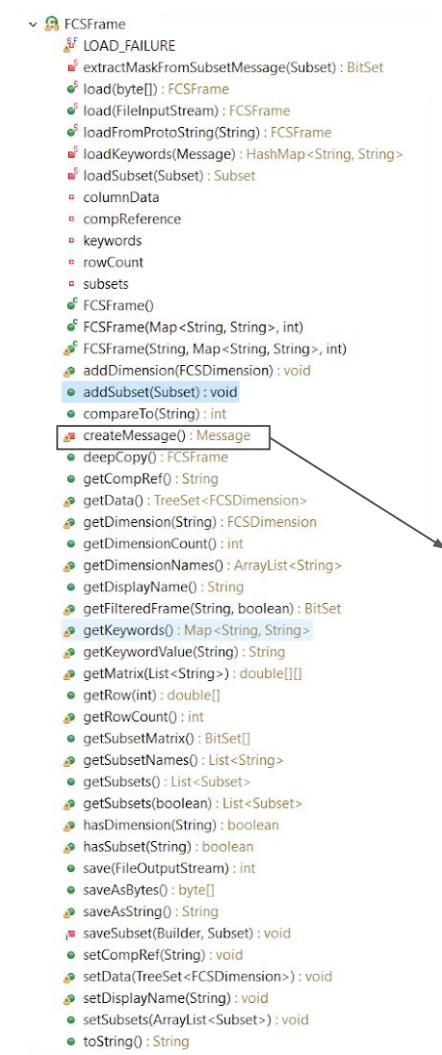
1:1 mapping to an FCS File

FCS Dimensions

Also stores gate membership

Serializes to disk or wire using protobuf.

Row ID	FCS Frame
Row 0	04EHV59_FLK 2008_A05.fcs 158340 FSC-A Min display: 0.0, Max display: 262144.0 FSC-H Min display: 0.0, Max display: 262144.0 SSC-A Min display: 0.0, Max display: 262144.0 SSC-H Min display: 0.0, Max display: 262144.0 CD154: FITC-A t=262144.0, w=0.5, m=4.5, a=0.0 MIP1b: PE-A t=262144.0, w=0.5, m=4.5, a=0.0 CD4: PerCP-Cy5-5-A t=262144.0, w=0.5, m=4.5, a=0.0 CD45RA: PE-Cy7-A t=262144.0, w=0.5, m=4.5, a=0.0 IFNg: Horizon V450-A t=262144.0, w=0.5, m=4.5, a=0.0 CD8 V500: AmCyan-A t=262144.0, w=0.5, m=4.5, a=0.0 Blue Vid-A t=262144.0, w=0.5, m=4.5, a=0.0 IL2: APC-A t=262144.0, w=0.5, m=4.5, a=0.0 TNFa: Alexa Fluor 700-A t=262144.0, w=0.5, m=4.5, a=0.0 CD3: APC-Cy7-A t=262144.0, w=0.5, m=4.5, a=0.0 Time Min display: 0.0, Max display: 262144.0
Row 1	04EHV59_NS_A01.fcs 164430 FSC-A Min display: 0.0, Max display: 262144.0 FSC-H Min display: 0.0, Max display: 262144.0 SSC-A Min display: 0.0, Max display: 262144.0 SSC-H Min display: 0.0, Max display: 262144.0 CD154: FITC-A t=262144.0, w=0.5, m=4.5, a=0.0 MIP1b: PE-A t=262144.0, w=0.5, m=4.5, a=0.0 CD4: PerCP-Cy5-5-A t=262144.0, w=0.5, m=4.5, a=0.0 CD45RA: PE-Cy7-A t=262144.0, w=0.5, m=4.5, a=0.0 IFNg: Horizon V450-A t=262144.0, w=0.5, m=4.5, a=0.0 CD8 V500: AmCyan-A t=262144.0, w=0.5, m=4.5, a=0.0



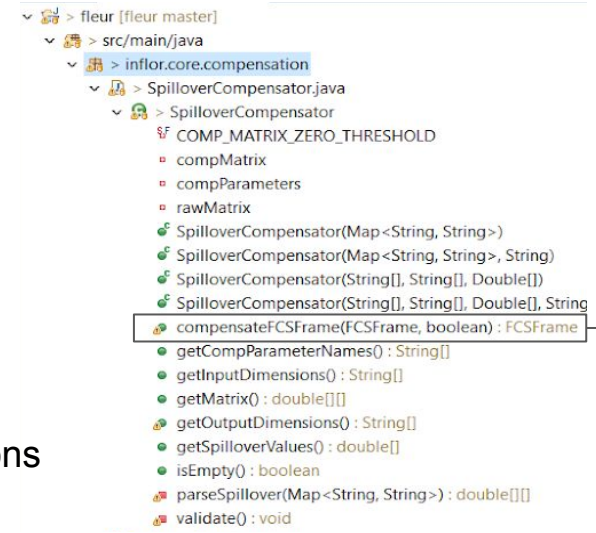

```
private Message createMessage() {
  // create the builder
  final Message.Builder messageBuilder = Message.newBuilder();
  messageBuilder.setId(this.getID());
  messageBuilder.setEventCount(this.rowCount);

  // add the dimension names.
  for (final String name : getDimensionNames()) {
    messageBuilder.addDimNames(name);
  }

  // add the keywords.
  for (Entry<String, String> s : keywords.entrySet()) {
    final String key = s.getKey();
    final String value = s.getValue();
    final Message.Keyword.Builder keyBuilder = Message.Keyword.newBuilder();
    keyBuilder.setKey(key);
    keyBuilder.setValue(value);
    final Message.Keyword newKeyword = keyBuilder.build();
    messageBuilder.addKeyword(newKeyword);
  }
  // add the FCS Dimensions.
  for (FCSDimension dim : columnData) {
    Message.Dimension.Builder dimBuilder = Message.Dimension.newBuilder();
    dimBuilder.setIndex(dim.getIndex());
    dimBuilder.setPnn(dim.getShortName());
    dimBuilder.setPns(dim.getStainName());
    dimBuilder.setPneF1(dim.getPNEF1());
    dimBuilder.setPneF2(dim.getPNEF2());
    dimBuilder.setPnr(dim.getRange());
    dimBuilder.setId(dim.getID());

    // Add the numeric data
    final double[] rawArray = dim.getData();
    for (final double value : rawArray) {
      dimBuilder.addData(value);
    }
    final Message.Dimension fcsdim = dimBuilder.build();
    messageBuilder.addDimension(fcsdim);
  }
```

Compensation

Parse comp matrix from the header of an FCS File

Apply the compensation

Uses ejml for matrix operations




```java
public FCSFrame compensateFCSFrame(FCSFrame dataFrame, boolean retainUncomped) throws InvalidProtocolBufferExcep
  FCSFrame newFrame = dataFrame.deepCopy();
  double[][] x = new double[compParameters.length][newFrame.getRowCount()];

  for (int i = 0; i < compParameters.length; i++) {
    Optional<FCSDimension> dimension = FCSUtilities.findCompatibleDimension(newFrame, compParameters[i]);
    if (!dimension.isPresent()) {
      throw new IllegalArgumentException("DataFrame does not contain matching parameters: " + compParameters[i])
    }
    x[i] = dimension.get().getData();
  }
  DenseMatrix64F xt = new DenseMatrix64F(x);
  CommonOps.transpose(xt);
  // not sure about mutability.
  mult(xt.copy(), compMatrix, xt);
  CommonOps.transpose(xt);

  x = new double[compParameters.length][newFrame.getRowCount()];
  for (int i = 0; i < compParameters.length; i++) {
    for (int j = 0; j < newFrame.getRowCount(); j++) {
      double newVal = xt.get(i, j);
      x[i][j] = newVal;
    }
  }
  for (int i = 0; i < compParameters.length; i++) {
    Optional<FCSDimension> dimension = FCSUtilities.findCompatibleDimension(newFrame, compParameters[i]);
    dimension.get().setData(x[i]);
    dimension.get().setShortName("[" + dimension.get().getShortName() + "]");
  }
  newFrame.setCompRef(this.getID());

  if (retainUncomped){
    for (int i = 0; i < compParameters.length; i++) {
      Optional<FCSDimension> dimension = FCSUtilities.findCompatibleDimension(dataFrame, compParameters[i]);
      newFrame.addDimension(dimension.get());
    }
  }

  return newFrame;
}
```

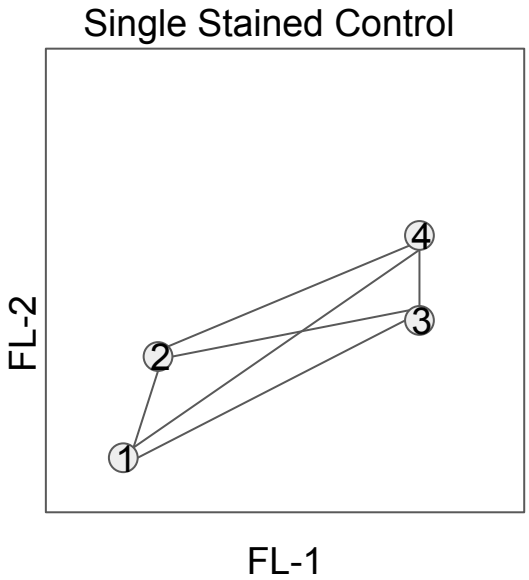
New approach:
Thiel-Sen Estimator for Compensation

Calculates compensation based on single stained controls without drawing gates.

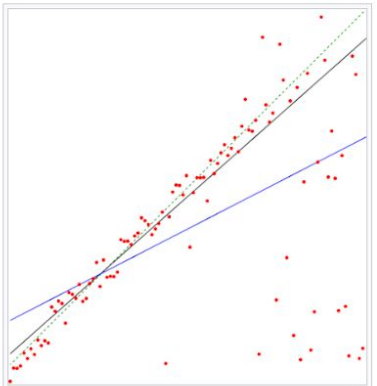
Algorithm takes the median of the pairwise slopes.

Significantly eases the automation of compensation.

Unpublished and needs more validation but appears to work very well.



Pairing	Slope
1-2	100
1-3	1
1-4	1.5
2-3	0.75
2-4	1
3-4	10000

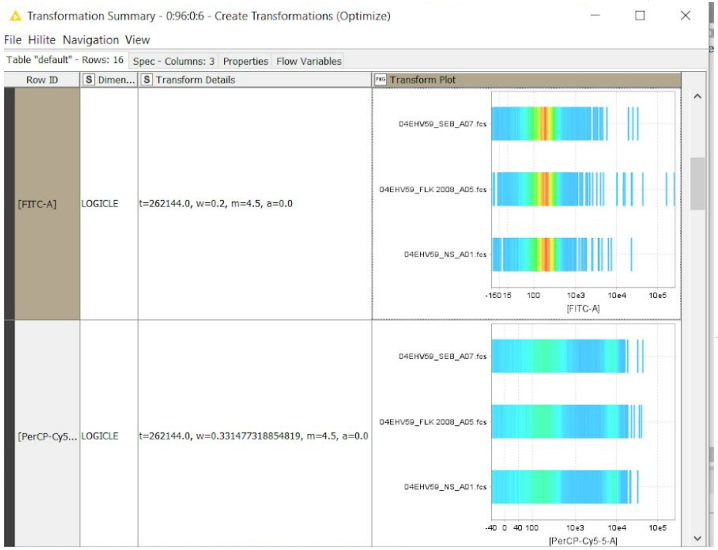


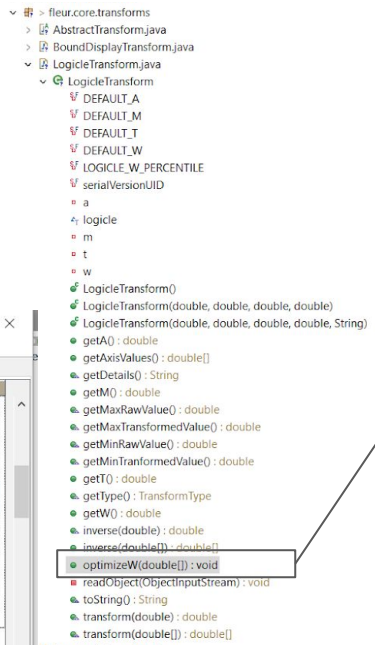
https://en.wikipedia.org/wiki/Theil%E2%80%93Sen_estimator#

Display Transformation

Supports logicle scaling (Stanford)

Including W estimation





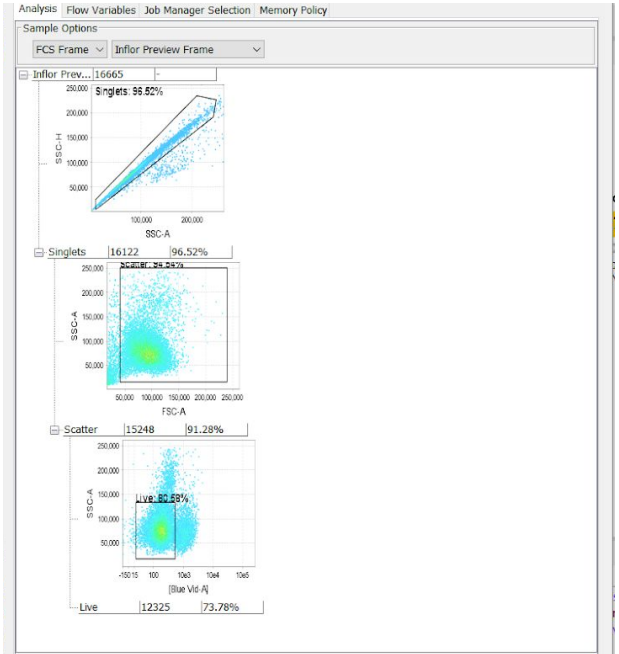

```java
public void optimizeW(double[] data) {
  /**
   * Based on the percentile method suggested by Parks/Moore.
   */
  double lowerBound = new Percentile().evaluate(data, LOGICLE_W_PERCENTILE);
  if (lowerBound < 0){
      this.w = (m - Math.log10(t / Math.abs(lowerBound))) / 2;
  } else {
      this.w = 0.2;
  }
  this.logicle = new FastLogicle(logicle.T, this.w, logicle.M, logicle.A);
}
```

Gating

Familiar hierarchical gate editor

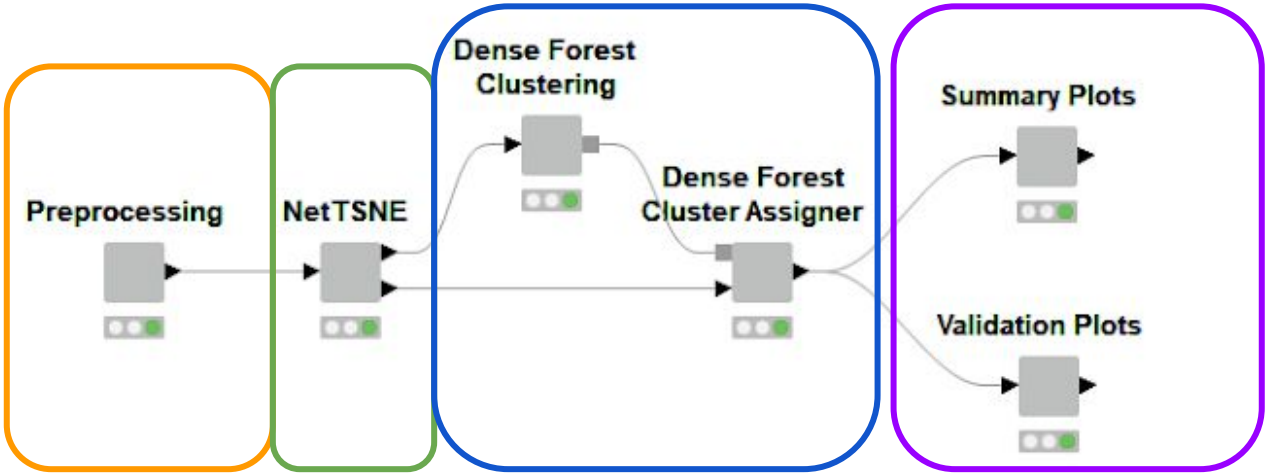
Limited but sufficient functionality

Gate membership is exportable!

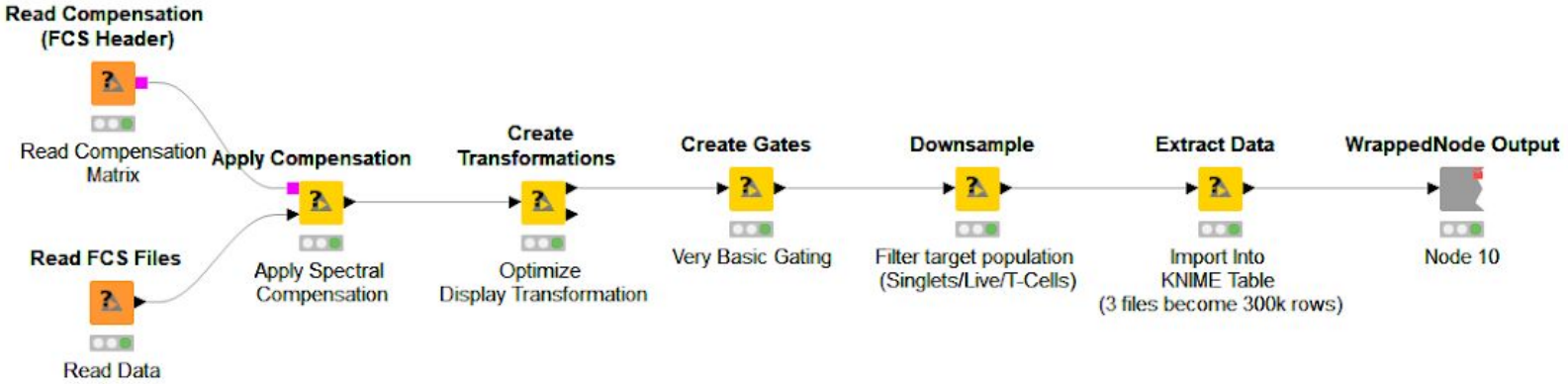



```java
@Override
public Component getTreeCellRendererComponent(JTree tree, Object value, boolean selected,
    boolean expanded, boolean leaf, int row, boolean hasFocus) {
  DefaultMutableTreeNode node = (DefaultMutableTreeNode) value;
  node.breadthFirstEnumeration();

  List<AbstractGate> gates = extractGates(node.getUserObjectPath());
  if (referenceData != null) {
    BitSet mask = GateUtilities.applyGatingPath(referenceData, gates, transforms);
    // node is root
    if (node.isRoot()) {
      String[][] tableRow = new String[][] {
          {referenceData.getDisplayName(), Integer.toString(referenceData.getRowCount()), "-"}};
      JTable table = new JTable(tableRow, columnNames);
      formatTable(selected, table);
      return table;
      // node is a gate
    } else if (node.getUserObject() instanceof AbstractGate) {
      AbstractGate gate = (AbstractGate) node.getUserObject();
      String[][] tableRow = new String[][] {{gate.getLabel(),
          Integer.toString(mask.cardinality()), BitSetUtils.frequencyOfParent(mask, 2)}};
      JTable table = new JTable(tableRow, columnNames);
      formatTable(selected, table);
      return table;
      // node is plot
    } else if (node.getUserObject() instanceof ChartSpec) {
      FCSFrame filteredFrame = FCSUtilities.filterFrame(mask, referenceData);
      ChartSpec spec = (ChartSpec) node.getUserObject();
      AbstractFCChart plot = PlotUtils.createPlot(spec);
      JFreeChart chart = plot.createChart(filteredFrame, transforms);
      formatChart(selected, expanded, leaf, chart);
      FCSChartPanel panel = new FCSChartPanel(chart, spec, filteredFrame, transforms);
      List<AbstractGate> siblingGates = findSiblingGates(node);
      siblingGates.stream().filter(gate -> ChartUtils.gateIsCompatibleWithChart(gate, spec))
          .map(ChartUtils::createAnnotation).forEach(panel::createGateAnnotation);
      panel.setPreferredSize(new Dimension(220, 200));
      return panel;
    }
  }
  return new JLabel("Unsupported node type.");
}
```

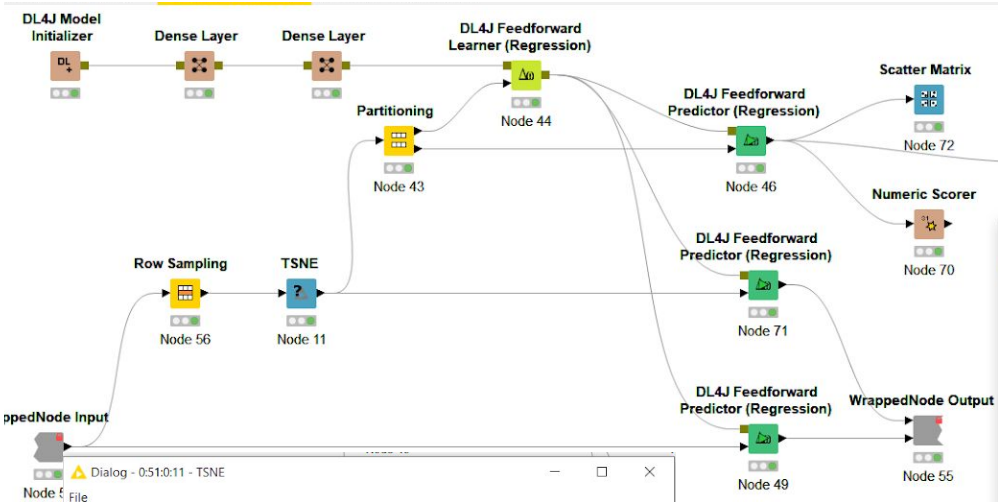
Use case: Exploratory Analysis



Preprocessing



NetTSNE


DL4J Model Initializer
Dense Layer
Dense Layer
DL4J Feedforward Learner (Regression)
Node 44
Partitioning
Node 43
DL4J Feedforward Predictor (Regression)
Node 46
Scatter Matrix
Node 72
Scatter Plot (JFreeChart)
Node 73
Numeric Scorer
Node 70
Row Sampling
Node 56
TSNE
Node 11
DL4J Feedforward Predictor (Regression)
Node 71
DL4J Feedforward Predictor (Regression)
Node 49
WrappedNode Output
Node 55
ppedNode Input

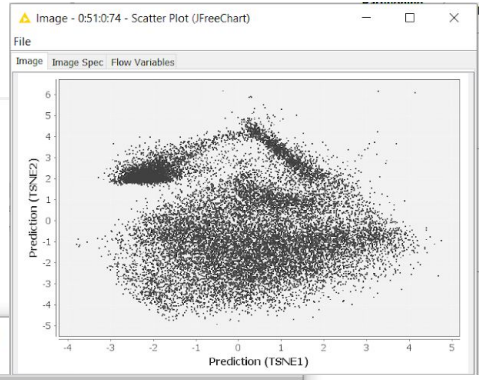
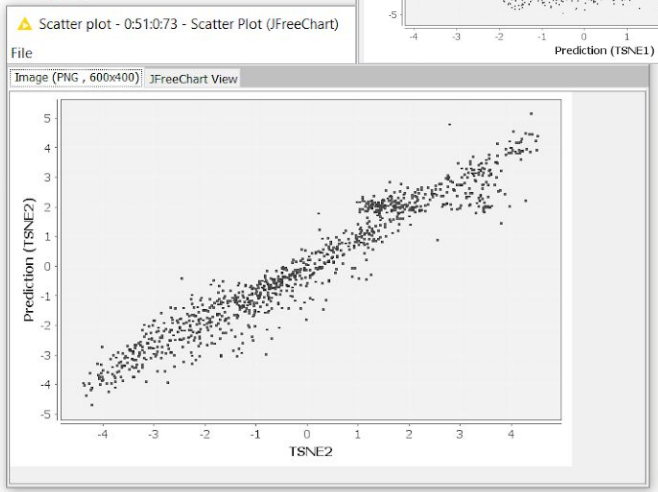
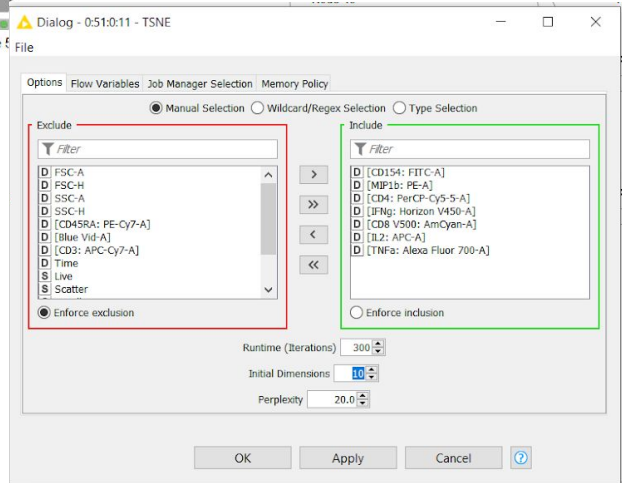
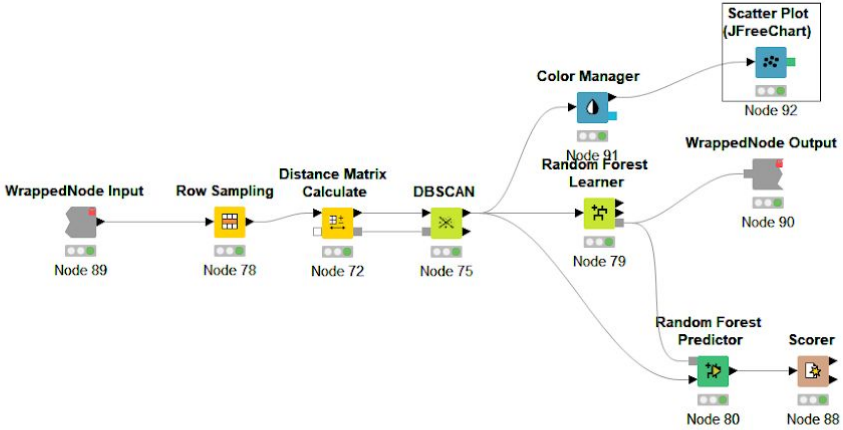


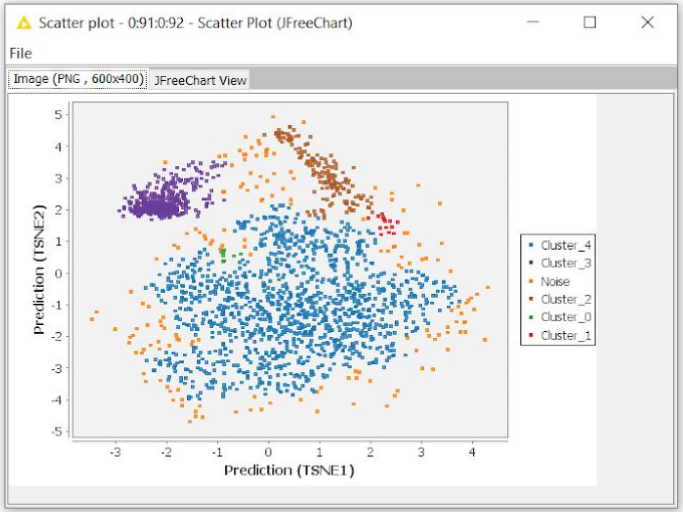
Image - 0:51:0:74 - Scatter Plot (JFreeChart)
File
Image
Image Spec
Flow Variables
Prediction (TSNE2)
Prediction (TSNE1)


Scatter plot - 0:51:0:73 - Scatter Plot (JFreeChart)
File
Image (PNG , 600x400)
JFreeChart View
Prediction (TSNE2)
TSNE2

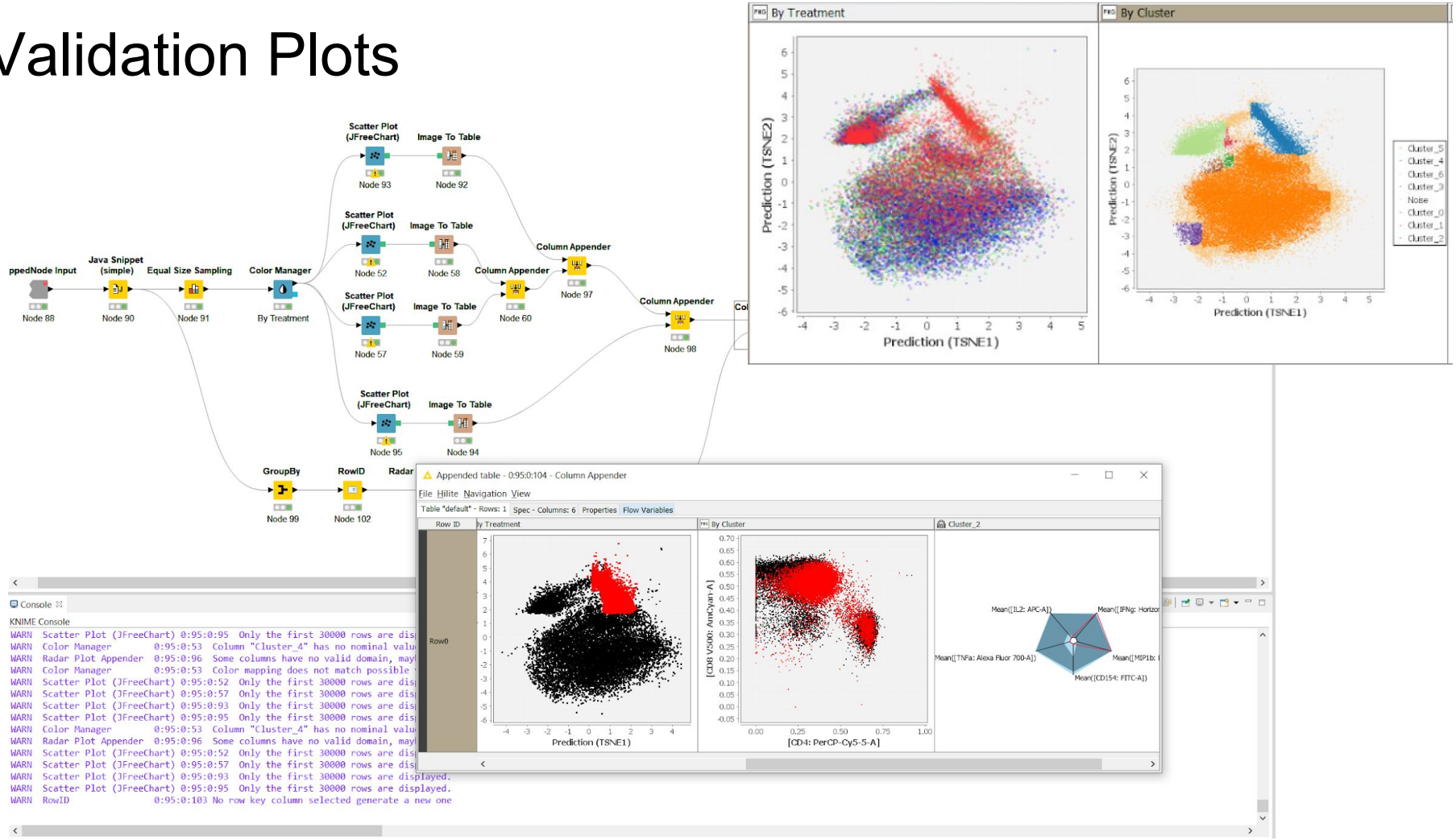

Dialog - 0:51:0:11 - TSNE
File
Options
Flow Variables
Job Manager Selection
Memory Policy
Manual Selection
Wildcard/Regex Selection
Type Selection
Exclude
Filter
D FSC-A
D FSC-H
D SSC-A
D SSC-H
D [CD45RA: PE-Cy7-A]
D [Blue Vid-A]
D [CD3: APC-Cy7-A]
D Time
S Live
S Scatter
Enforce exclusion
Include
Filter
D [CD154: FITC-A]
D [MIP1b: PE-A]
D [CD4: PerCP-Cy5-5-A]
D [IFNg: Horizon V450-A]
D [CD8 V500: AmCyan-A]
D [IL2: APC-A]
D [TNFa: Alexa Fluor 700-A]
Enforce inclusion
Runtime (Iterations)
300
Initial Dimensions
10
Perplexity
20.0
OK
Apply
Cancel

Dense Forest Clustering

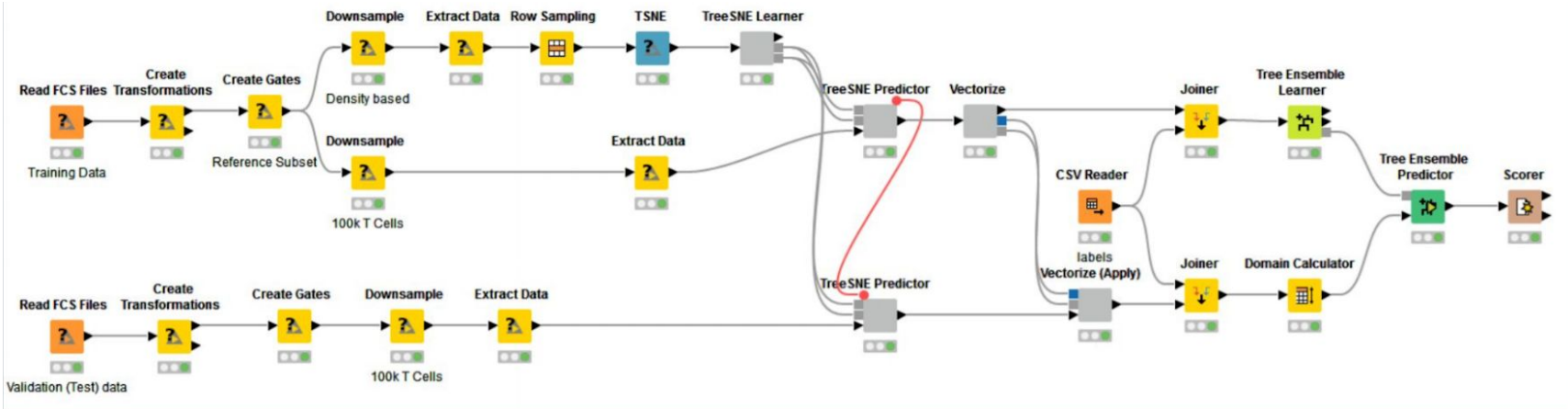




Validation Plots



Use case: Outcome driven modeling.



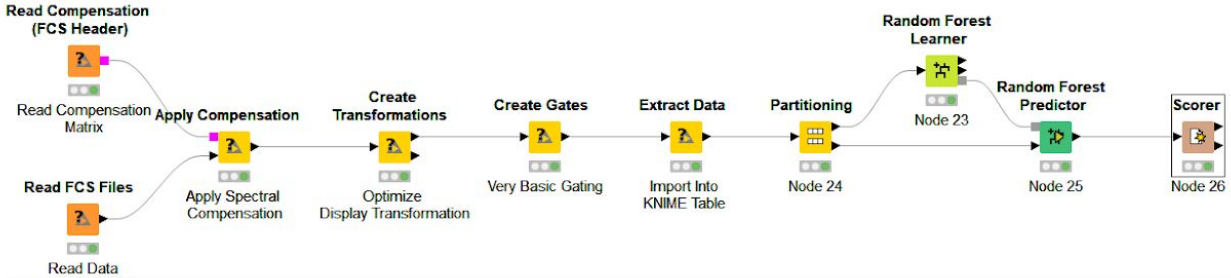
Sample Tre...	negctrl	sebctrl
negctrl	26	0
sebctrl	0	10

Correct classified: 36 Wrong classified: 0

Accuracy: 100 % Error: 0 %

Cohen's kappa (κ) 1

Use case: Automation of routine gating.



Confusion Matrix - 0:26 - Scorer

File Hilite

SubsetN...	Ungated\...	Ungated\...	Ungated\...	Ungated\...	Ungated\...	Ungated	Ungated\...	Ungated\...	Ungated\...	Ungated\...	Ungated\...	Ungated\...	Ungated\...	Ungated\...	Ungated\...	Ungated\...
Ungated\L...	30781	21	0	0	0	0	0	0	0	2	0	0	0	0	0	0
Ungated\L...	78	5662	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ungated\S...	0	0	1131	0	1	3	2	8	0	0	0	0	0	0	0	0
Ungated\S...	1	0	0	5399	9	0	0	0	0	0	1	1	0	0	0	0
Ungated\S...	0	0	0	56	3516	0	0	0	0	0	8	0	0	0	0	0
Ungated	0	0	27	0	0	661	0	1	0	0	2	0	0	0	0	0
Ungated\L...	0	0	0	0	0	0	1109	0	0	0	0	0	6	0	0	0
Ungated\T...	0	0	3	0	0	0	0	311	0	0	0	0	1	0	0	0
Ungated\L...	3	18	0	0	0	0	0	0	14	2	0	0	0	0	0	0
Ungated\L...	78	0	0	0	0	0	0	0	0	80	0	0	0	0	0	0
Ungated\S...	0	0	0	3	83	0	0	0	0	0	501	6	0	0	0	0
Ungated\S...	0	0	0	31	0	0	0	0	0	0	8	141	0	0	0	0
Ungated\L...	0	0	0	0	0	0	5	0	0	0	0	0	135	0	0	0
Ungated\T...	0	0	0	0	0	0	0	2	0	0	0	0	0	5	0	0
Ungated\L...	0	0	0	0	0	0	0	0	0	0	0	0	5	0	4	0
Ungated\L...	0	0	0	0	0	0	7	0	0	0	0	0	2	0	0	1

Correct classified: 49,451 | Wrong classified: 484

Accuracy: 99.031 % | Error: 0.969 %

Cohen's kappa (κ) 0.983

WARN BlobWrapperDataCell Blob file location "C:\Windows\System32\col_0\000\000" does not exist

End

Contact
aaron.hart@gmail.com

Fleur
https://github.com/Landysh/fleur

KNIME
https://www.knime.com/

t-SNE
https://lvdmaaten.github.io/tsne/

Logicle Transform
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4761345/