Tuning Predictive Intelligence solutions: Classification

What's in this Success Insight

You'll learn to decide whether your classification model needs tuning, and if so, how to understand and tune your model for improved output.

This Success Insight provides best practices that address the following key questions:

- What does it mean to tune a classification model?
- How do I know if I need to tune my classification model?
- How do I improve precision and coverage for a classification model?

For prerequisites use the Success Workbook: Predictive Intelligence Readiness.

servicenow

Key insights

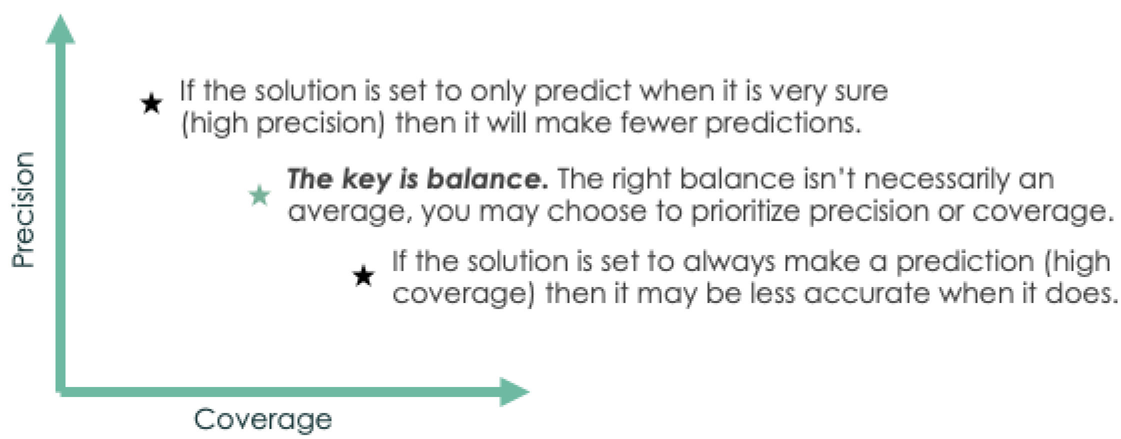
1. What does it mean to tune a clasification model?

Tuning a classification model means finding the balance between precision and coverage that will maximize the value you're receiving from Predictive Intelligence's classification capabilities.

Precision
Precision is the % of predicted records that are correct.

Coverage
Coverage is the % of records that are predicted—in other words, how much is automated.

The precision and coverage numbers are inversely related to each other:



If you're using classification solutions to detect incoming emails as a phishing scam or similar uses, there is an additional feature called recall that's useful.

Note: Machine learning solutions will rarely be 100% precise. The goal is for the machine learning model to be correct more than human agents. Models with a precision higher than 70% and coverage higher than 80% are typically more effective than human agents doing the same work.

2. How do I know if I need to tune my classification model?

To determine if tuning is needed, start by opening a solution record and click **Solution Visualization**. This opens a bubble chart with the information to assess your model.

Figure 1 is an example bubble chart with a good classification model. The blue circles represent the assignments groups. A larger circle size represents a larger distribution. For example, the "Deskside Support" assignment group represents 22% of the incidents from the training set and has a large distribution. What makes this a good model is that nearly all the blue bubbles (assignments groups) are in the upper far-right quadrant, which means they have good precision and good coverage. Note: The x-axis changes to "recall" in Orlando and beyond.

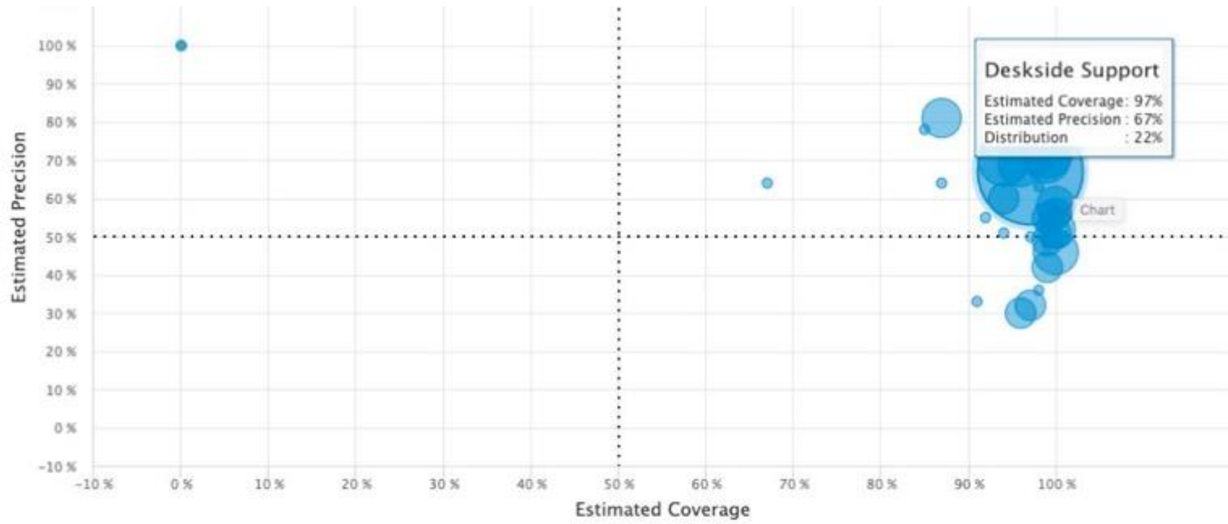


Figure 1: A good classification model

Figure 2 highlights outliers outside of the upper right quadrant. Those assignment groups should be excluded from the model because they have a small distribution and are rare cases where a human is likely required to determine when an incident goes to that assignment group. Our Predictive Intelligence documentation lists the steps to exclude classes from a model.

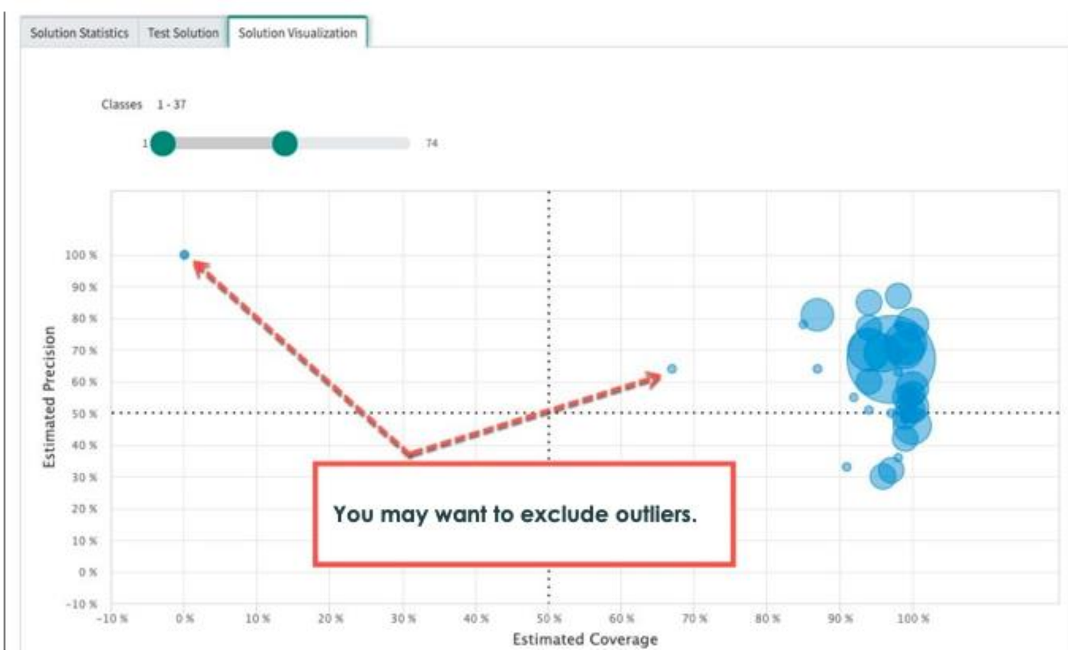


Figure 2: Assignment groups that should be excluded

Figure 3 represents a model that needs more adjustment. Many of the assignment groups with large distributions are scattered all over the graph, including several in the low-precision/low-coverage section (lower left quadrant).

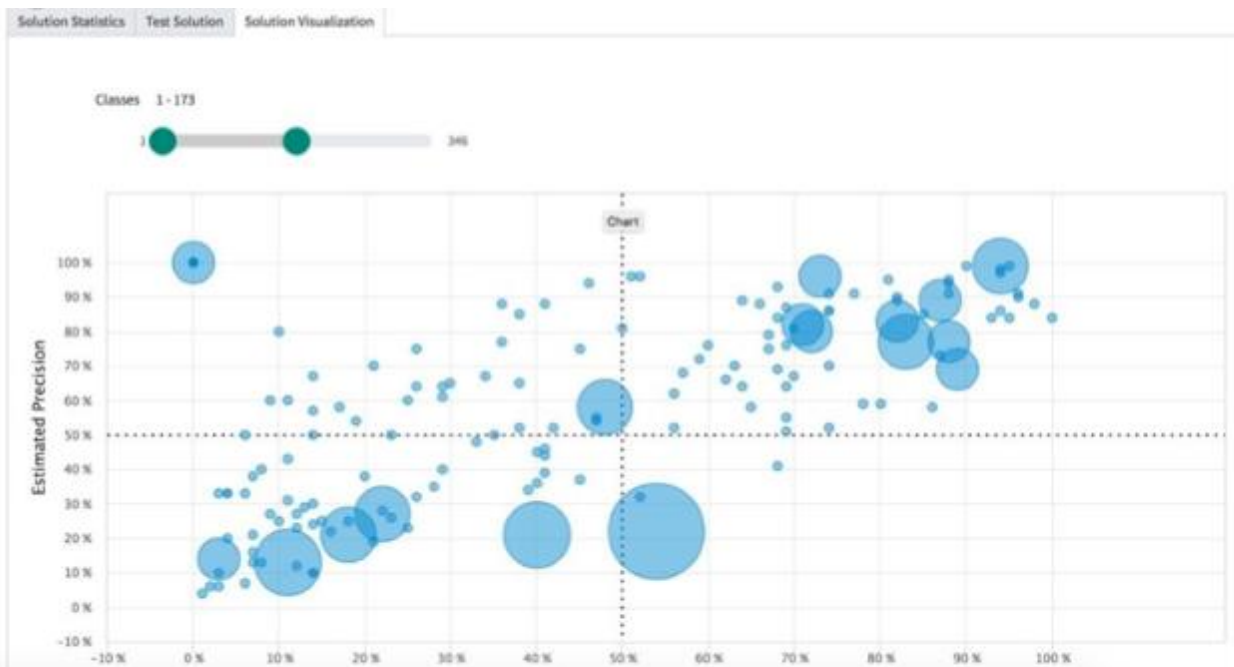


Figure 3: A model that needs adjustment

3. How do I improve precision and coverage for a classification model?

If your model looks closer to Figure 3 than Figure 1, set the target precision and coverage for a solution definition (available in Orlando) as shown in Figure 4. Setting the target level precision will allow Predictive Intelligence to automatically change the precision and coverage values to get close to the target precision. When choosing a target precision, remember that higher precision usually means lower coverage. If you choose 100% precision, it's likely to make few predictions and may not be as useful as a lower-precision target with a higher number of predictions.

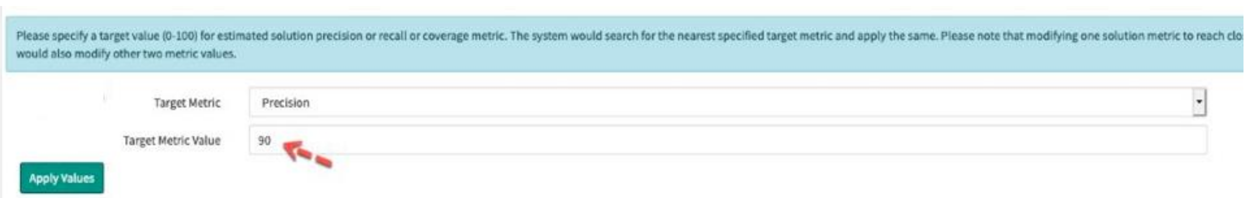


Figure 4: Where to set target precision

If, after setting the precision target, you're still not seeing most or all blue bubbles in the upper-right quadrant, there are additional methods to tune your model.

First, review your solution records. Below is an example of a solution record. There are multiple values for the output fields—also called class values—listed under **Class Confidence** and **Names**. In the example below, the class value **database** has an estimated precision of 89% and coverage of 100%. This means that out of 100 incidents predicted, about 89% will be correct and about 11% will be incorrect. It’s sometimes desirable to reduce the precision to increase the coverage or vice versa.

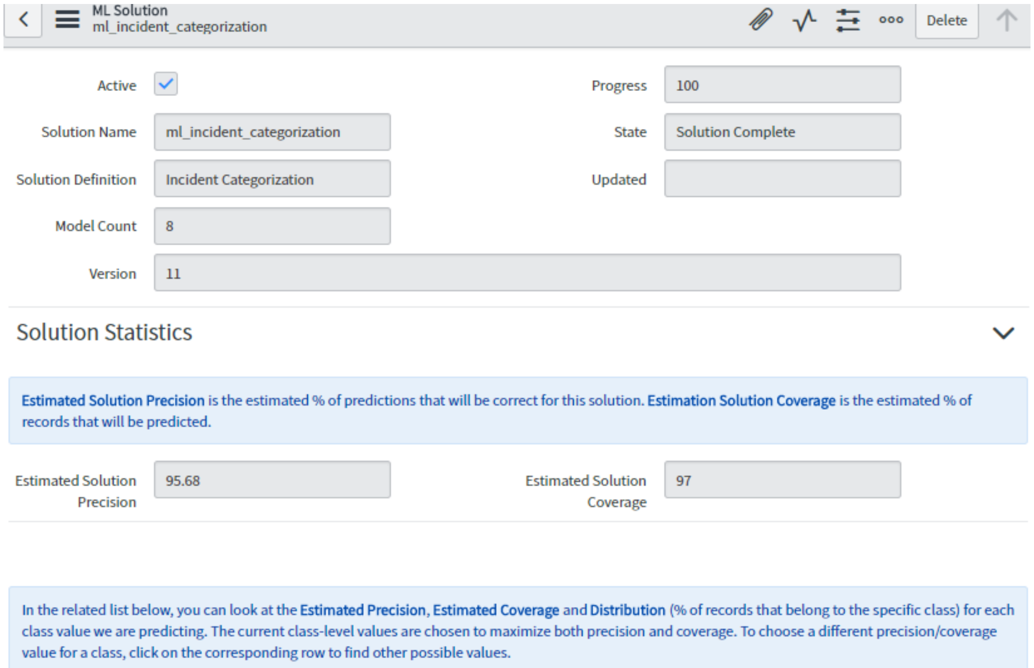
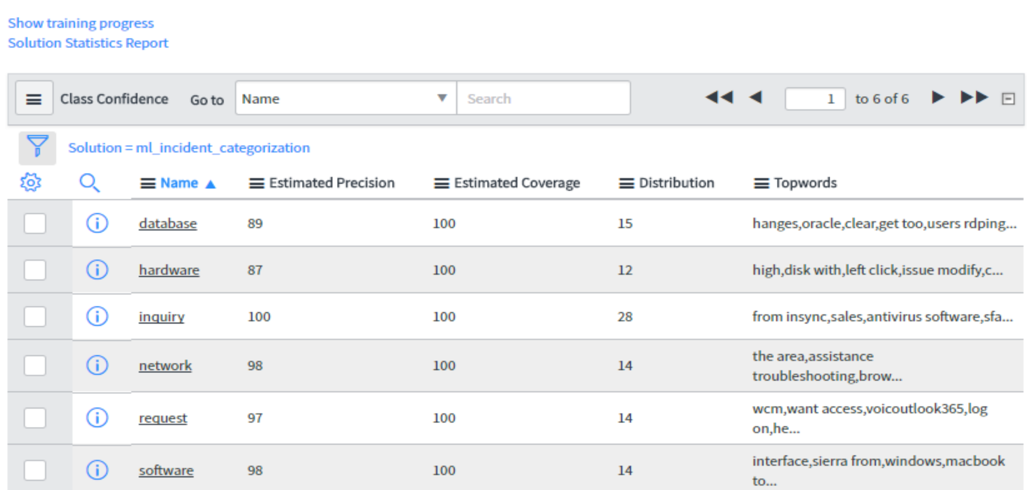


Figure 5: An example solution record



Name ▲	Estimated Precision	Estimated Coverage	Distribution	Topwords
database	89	100	15	hanges,oracle,clear,get too,users rdping...
hardware	87	100	12	high,disk with,left click,issue modify,c...
inquiry	100	100	28	from insync,sales,antivirus software,sfa...
network	98	100	14	the area,assistance troubleshooting,brow...
request	97	100	14	wcm,want access,voicoutlook365,log on,he...
software	98	100	14	interface,sierra from,windows,macbook to...

Figure 6: An example soution record

In the screenshot above you’ll find **software** located five lines below **database**. Below is a screen shot of what you see after you click **software**.

The default values for precision, coverage, and threshold are typically the optimal combination but can be altered.

servicenow

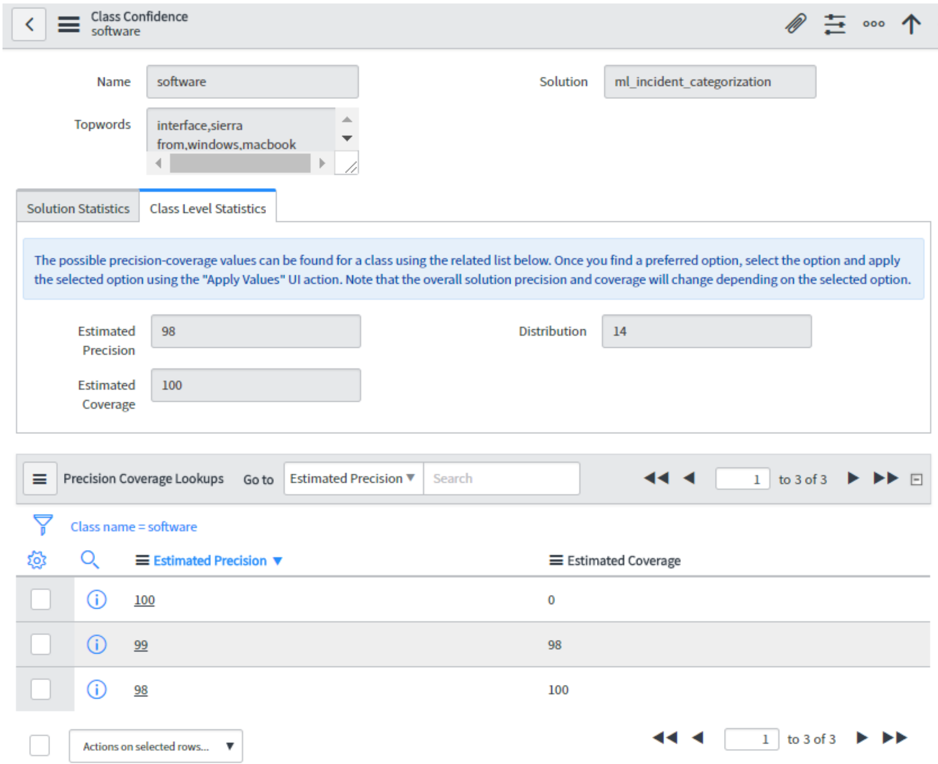


Figure 7: An example record with class level statistics

Once you review your solution record take these actions if you want to improve your model:

1. Check your data quality. Do you have assignment groups that mean the same thing with different spellings or empty/null assignment groups?

   If so, exclude those from the training set or clean up your data. An example of cleaning up data would be standardizing assignment group names and/or filling in empty or null assignment groups.

2. If you aren't sure what data to exclude and you have a large data set (200,000+ records), try a smaller data set. Try training the minimum 30,000 records, starting with the most recent, to see if it increases the precision. If it does, incrementally increase the data volume to see when precision starts declining and find the amount of data that optimizes precision and coverage.

3. Make sure the assignment groups are well distributed. For example, if more than 60% of the incidents fall under one assignment group (in the bubble chart example, the assignment group is Deskside Support), then you don't need a machine learning model to tell you where new incidents will land.

4. You don’t want to use inputs with many unique values. For example, a configuration item (CI) that may have millions or more records may make it harder for the machine learning algorithm to predict if you use CI as an input field.

5. Look at individual class level settings and see if a higher precision can be selected for that class.

6. If you started with the OOTB classification solution definition and it’s giving a low precision/coverage, try analyzing the incident/case data (it will be incident or case data depending on whether you are using Predictive Intelligene for ITSM, CSM, or HRSD) to determine what other inputs may help the system predict a more precise result. Running lists/reports with different input combinations should indicate whether additional inputs such as category, subcategory, or location could help improve the precision/coverage of the model.

7. Simplify your model. Don’t add many inputs to the solution definition without understanding the impact they will make—more is not necessarily better. If you have many inputs and your model has low precision/coverage, then delete all the inputs and start again with just a short description. Update and retrain, then add a few inputs and so forth until you get the desired precision.

8. It’s possible you won’t be able to predict assignment groups with just one model. Plan to use multiple models using different solutions definitions in each. In Figure 3, four classification models are used to predict the assignment groups.

The takeaway

Gaining confidence and skills for tuning classification models will take some experimentation! Each data set is different and your preferences for the precision and coverage will vary based on how you want to use the data. Before making any adjustments to your data or model, use tools available to you such as the bubble chart and reporting features to understand your solution set, and use the preset precision feature (available in Orlando). There are many contributing factors that impact your model, but as long as you understand your data and how to make adjustments to your model (tune it), you'll be equipped to get maximum value from Predictive Intelligence.

Additional resources

- Analytics, Intelligence and Reporting Community Forum

You'll find more resources on our Customer Success Center.

If you have any questions on this topic or you would like to be a contributor to future ServiceNow best practice content, please contact us at best.practices@servicenow.com.

servicenow

Customer Success Best Practices

ServiceNow's Best Practice Center of Excellence provides prescriptive, actionable advice to help you maximize the value of your ServiceNow investment.

Definitive guidance on a breadth of topics

Strategic

Management

Technical

Tactical

- Critical processes
- Expert insights
- Common pitfalls and challenges

Designed for:

- Executive sponsors
- Platform owners and teams
- Service and process owners

Created and vetted by experts

Best practice insights from customers, partners, and ServiceNow teams

Based on **thousands of successful implementations** across the globe

Distilled through a **rigorous process** to enhance your success

Proven to help you transform with confidence

Practical

Actionable

Value-added

Expert-validated

Get started today.

Visit Customer Success Center.

Contact your ServiceNow team for personalized assistance.