

ORACLE®

Oracle Digital Assistant

The Complete Training

Training your model

Safe Harbor Statement

The following is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, and timing of any features or functionality described for Oracle's products remains at the sole discretion of Oracle.

Topic agenda

- 1 ➤ Recap – the models in Oracle Digital Assistant
- 2 ➤ General guidelines for training your skills
- 3 ➤ Specific tips for training your skills
- 4 ➤ Training the digital assistant
- 5 ➤ Quality reports

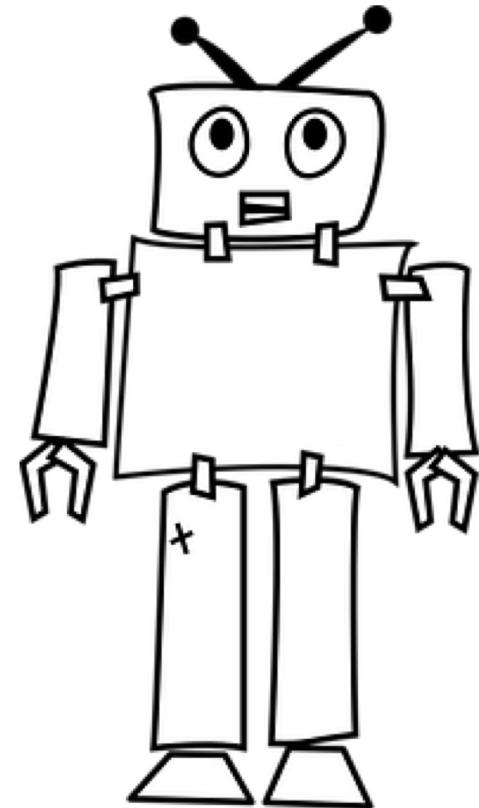
Topic agenda

- 1 ➤ Recap – the models in Oracle Digital Assistant
- 2 ➤ General guidelines for training your skills
- 3 ➤ Specific tips for training your skills
- 4 ➤ Training the digital assistant
- 5 ➤ Quality reports

Why have different training models?

- Trainer Ht
 - Rules based
 - Fast & best suited for small set of utterances
 - Good for new development
- Trainer Tm
 - Machine learning
 - Thrives on more and more data
 - Higher accuracy (especially data outside your utterances) if enough data
 - Already trained on “knowledge” of English language (NLP)
- Q&A

Your digital assistant is only as
smart as the data you train it with
(rubbish in, rubbish out)

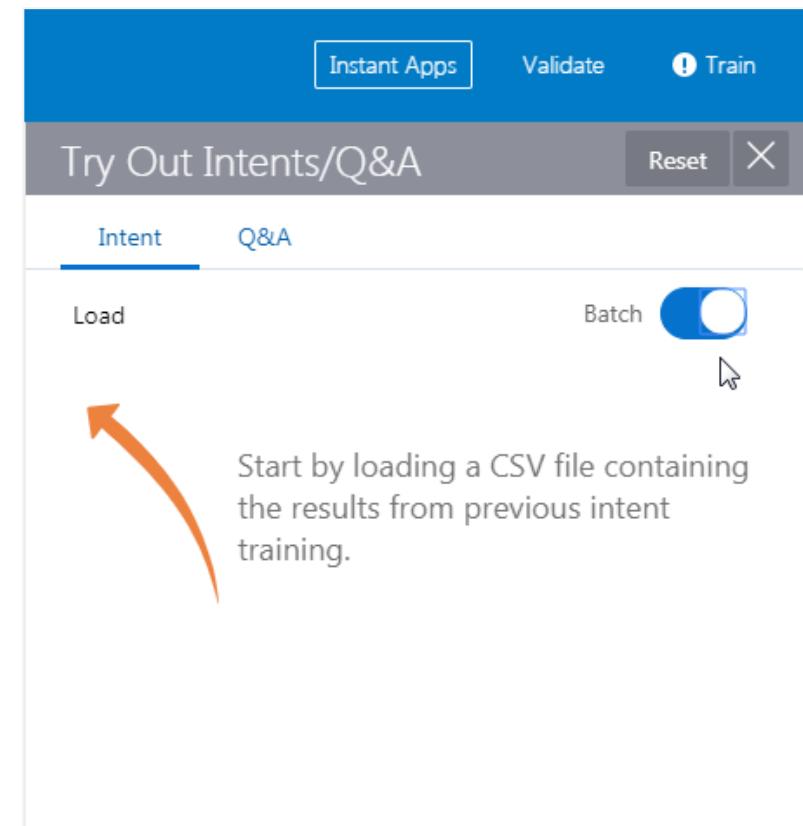


Topic agenda

- 1** → Recap – the models in Oracle Digital Assistant
- 2** → General guidelines for training your skills
- 3** → Specific tips for training your skills
- 4** → Training the digital assistant
- 5** → Quality reports

General guidelines for training your skills

- Your first release will have the language understanding of a 2 year old (if you are lucky!)
 - But that's ok, you need to start somewhere
- Plan for NLP improvement
 - Train, test, repeat
 - Use batch training feature
 - For all new data perform a 80/20 split
 - 80% training data, 20% testing data



General guidelines for training your skills

- Your goal is to train the model on real-world sample utterances
 - Synthesized utterances are no substitute for real world utterances
- Real-world data will likely need to be manually classified
 - It is not always clear what phrases map to an intent
 - Needs to be decided by project team: business, conversational designer, developer
- You shouldn't necessarily sanitize user data
 - Include common mis-spellings, slang, synonyms, abbreviations where it makes sense

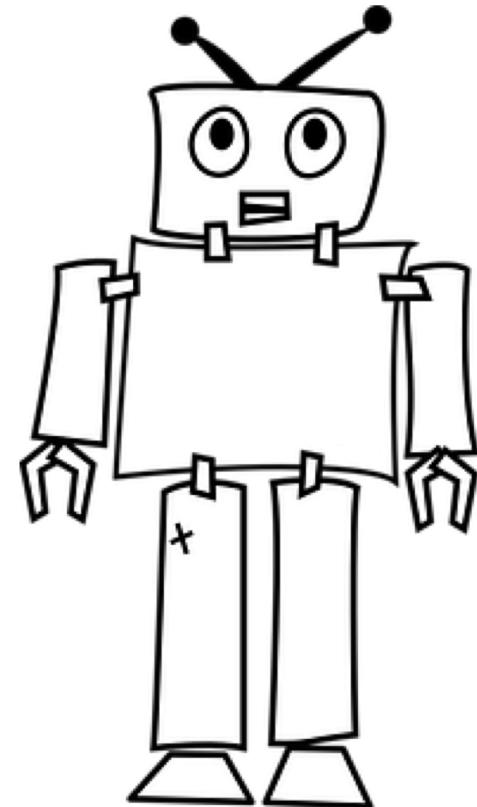
General guidelines for training your skills

- The golden rules to help you develop good understanding
 - For production use Trainer Tm
 - Real world utterances
 - Well classified
 - Train unresolved (anti-usecase) utterances
 - As many as you need to get the results you desire

Synthesizing utterances when you have no existing data

- Trainer Ht best place to start off
 - Generally better with smaller data sets
 - Plan to move to Trainer Tm as you gather more sample user inputs
- You have to synthesize sample utterance
 - Your primary goal is to help disambiguate intents
- A model has no inherent knowledge of what an utterance actually means
 - Frequency of words, sentence patterns, some knowledge of parts of speech, train with synonyms
- Train, test, repeat

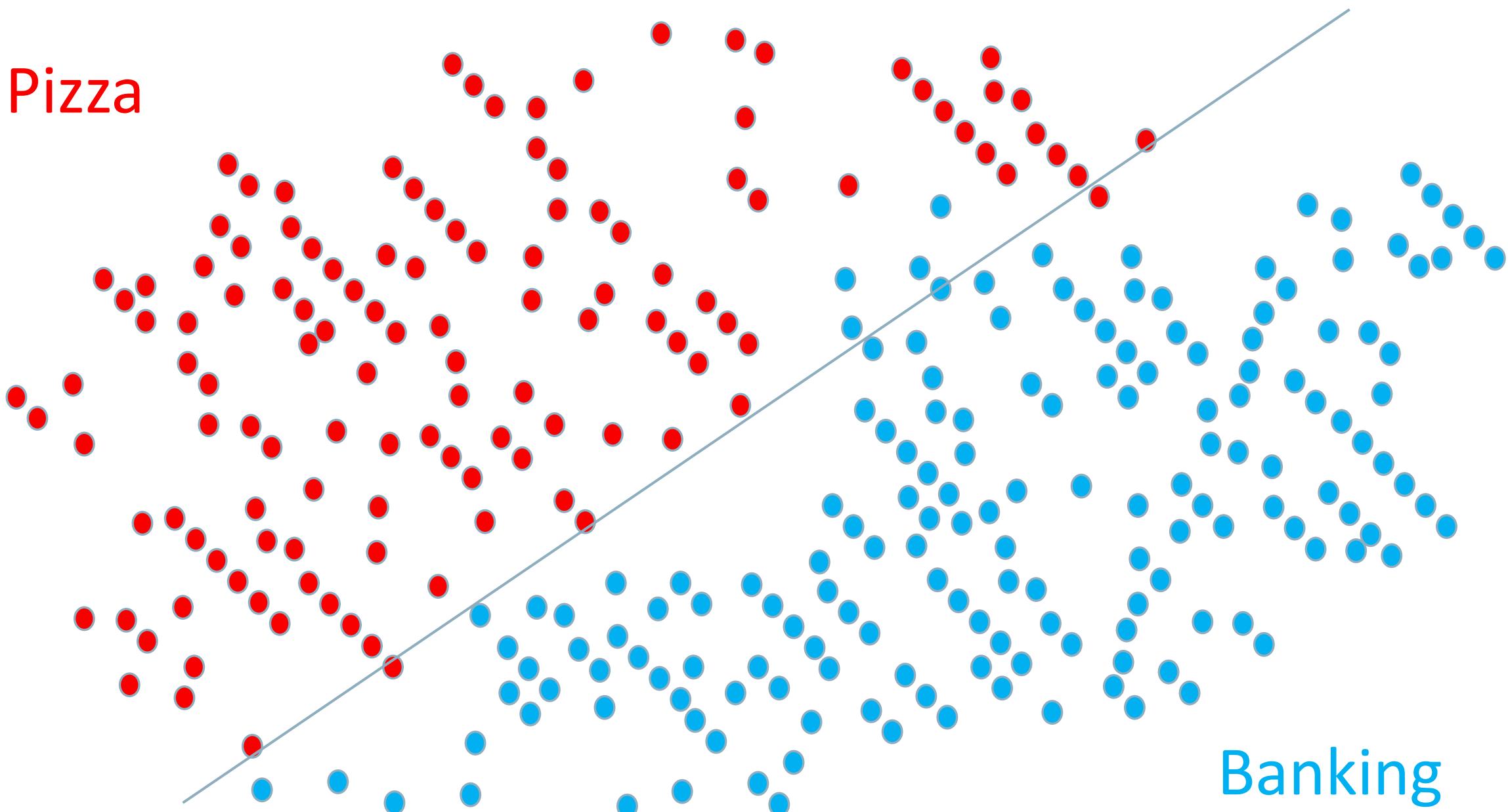
So if I am going to build a **good model** shouldn't I know how it works?



Pizza

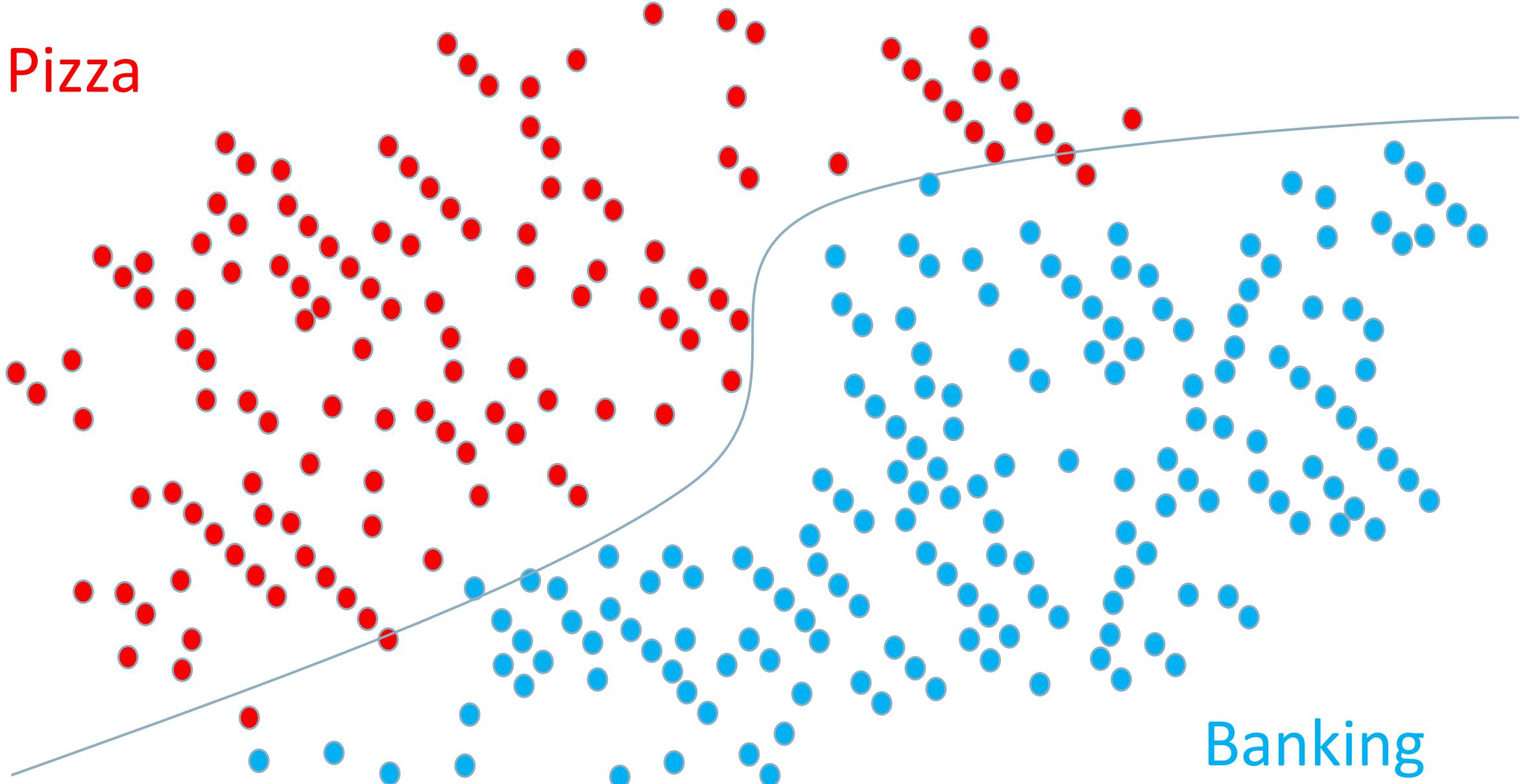


Pizza

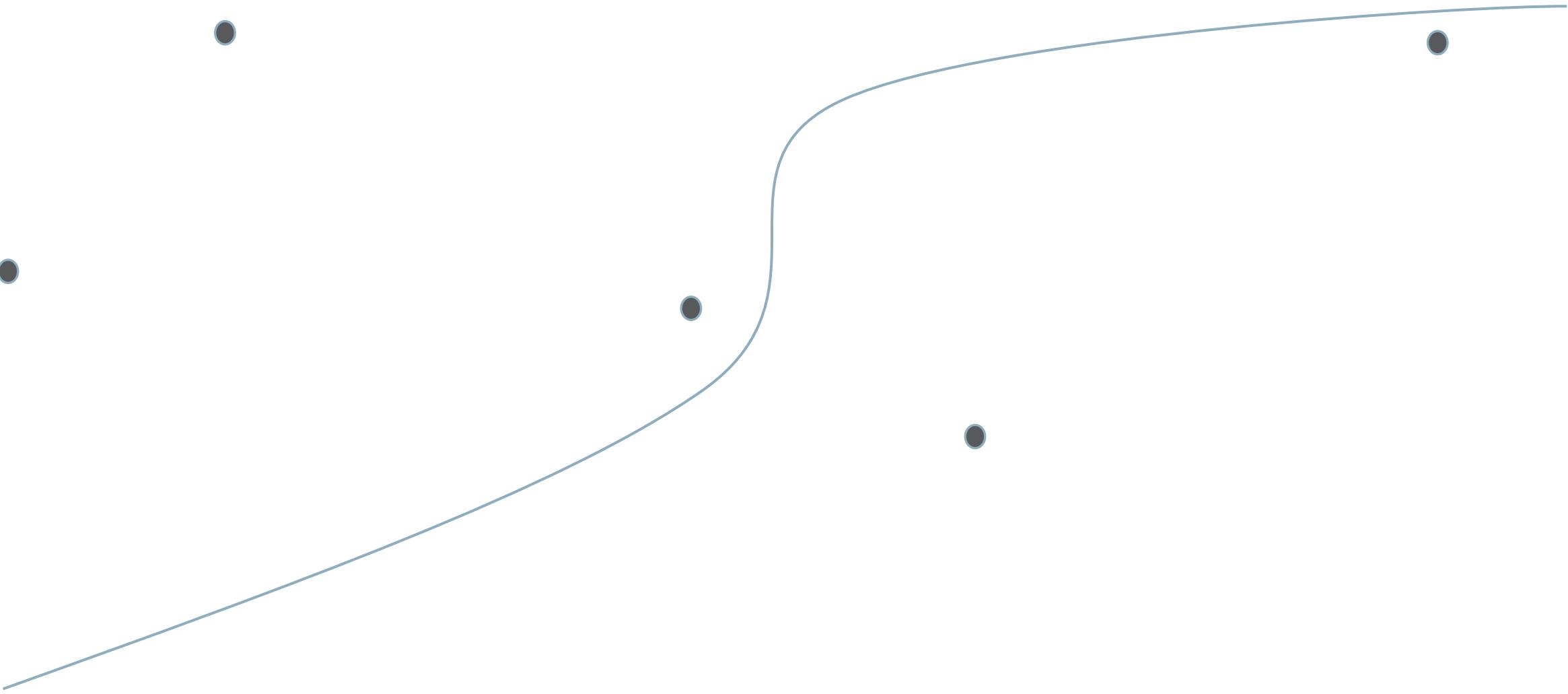


Banking

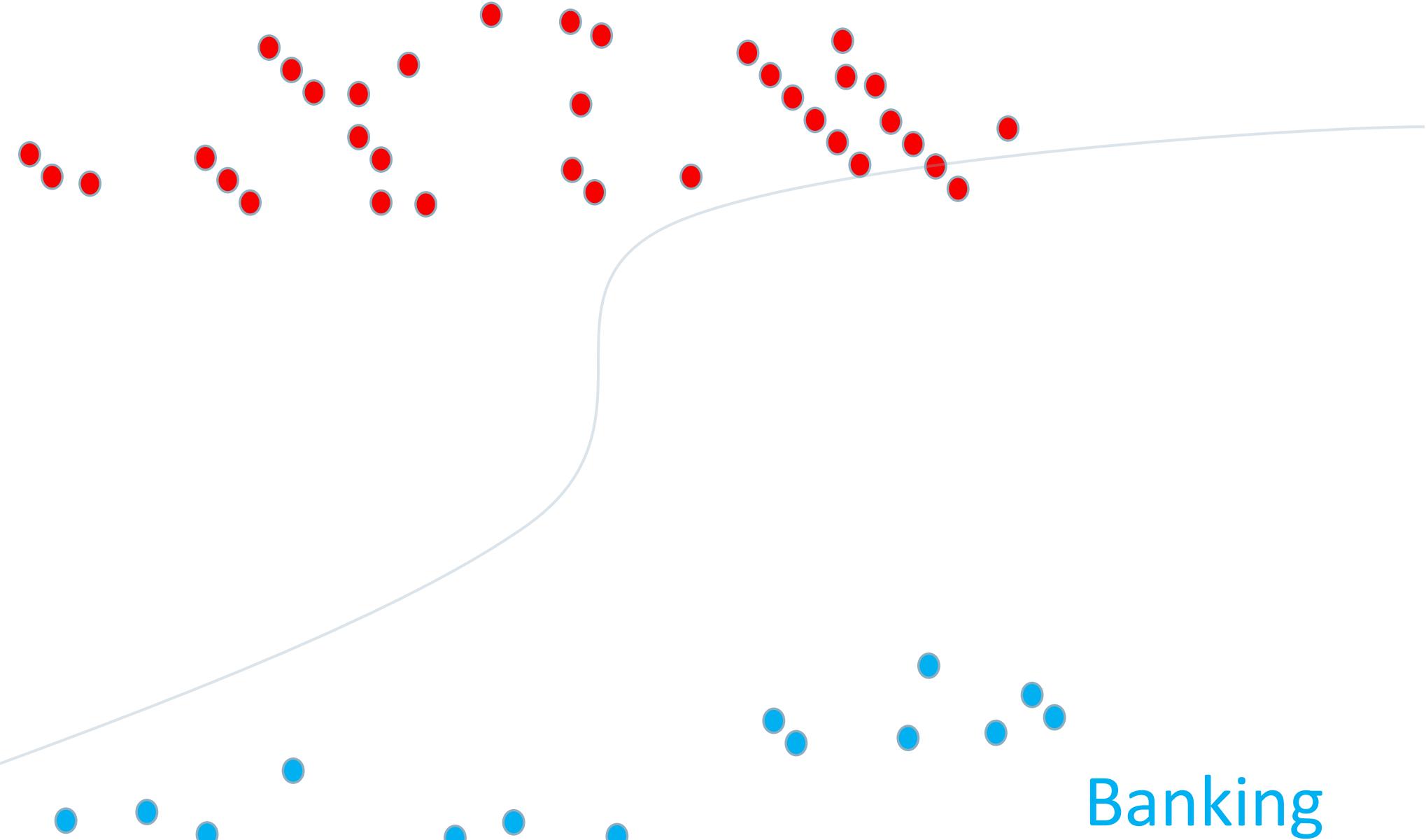
Pizza



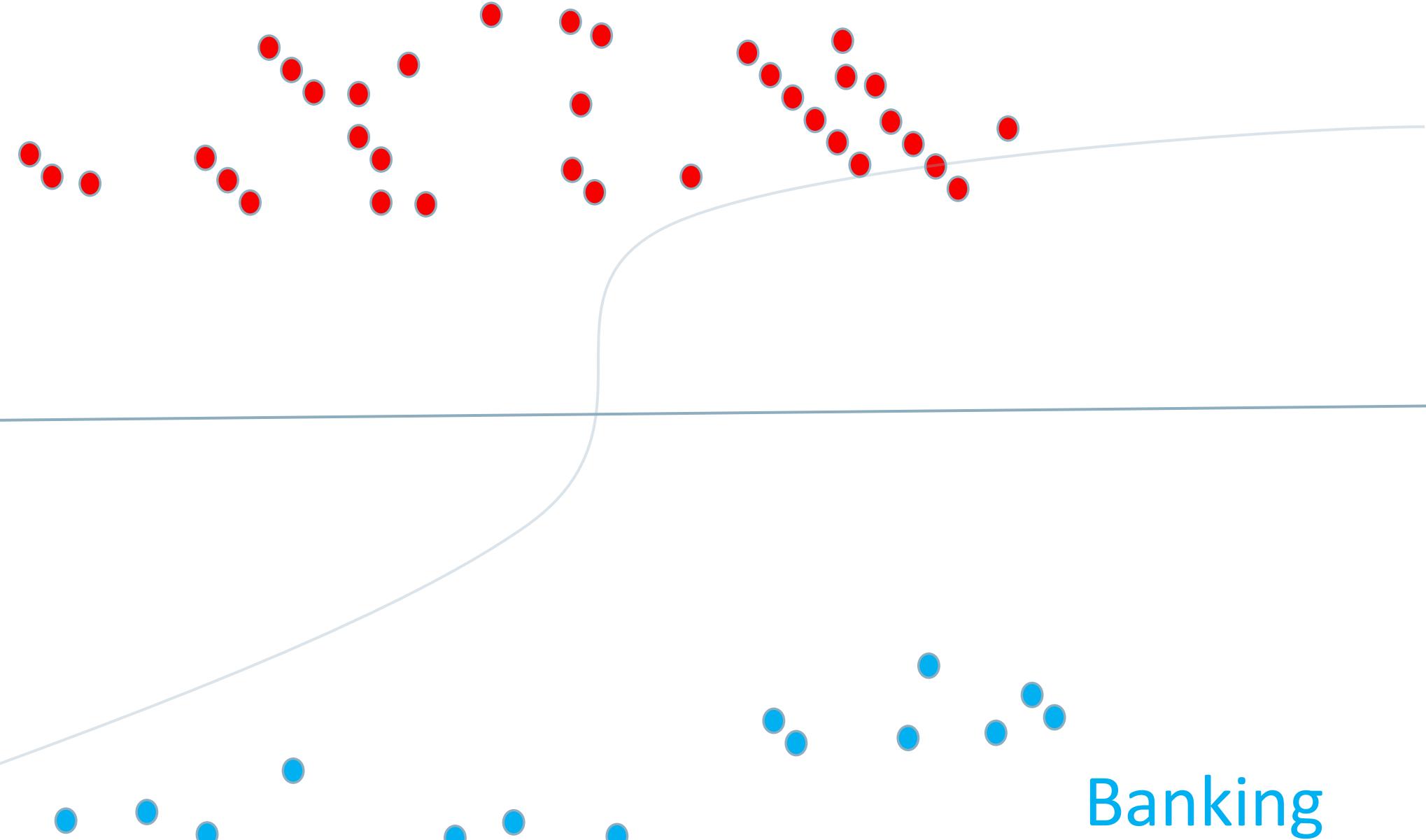
Pizza



Pizza

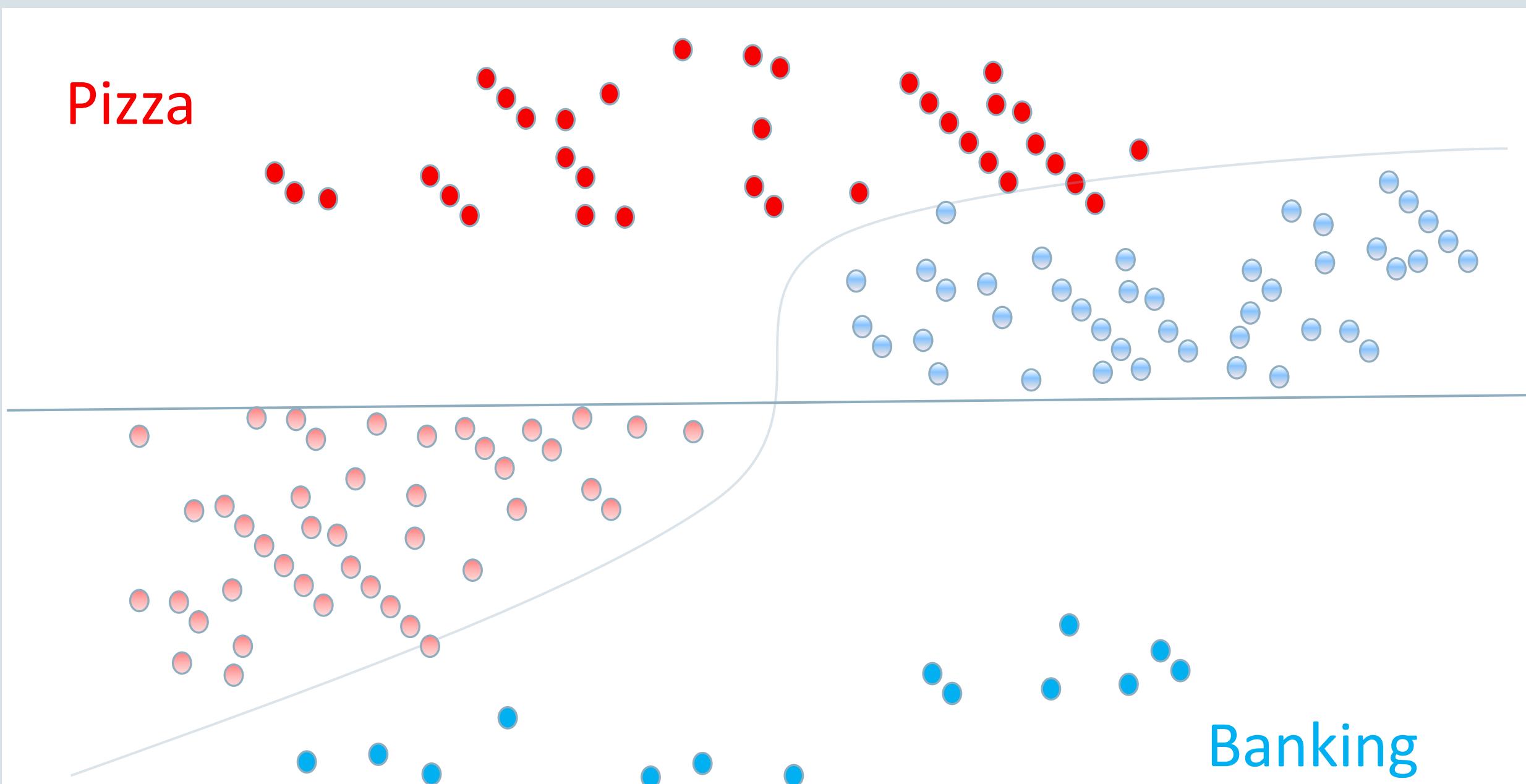


Pizza

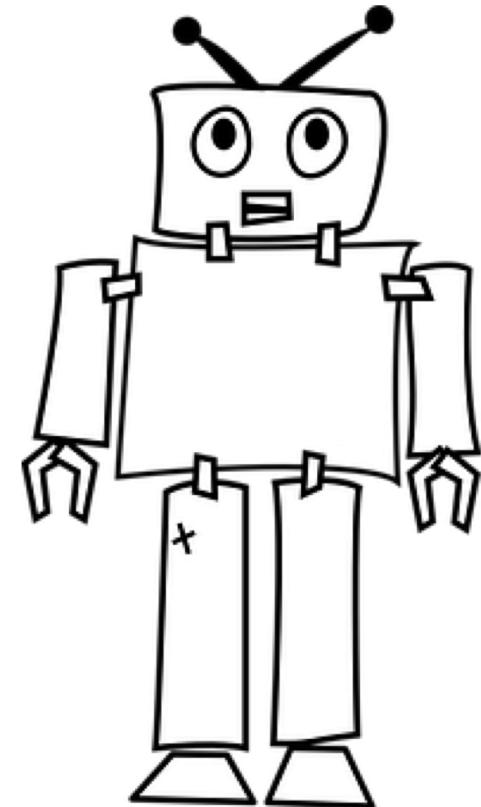


Pizza

Banking

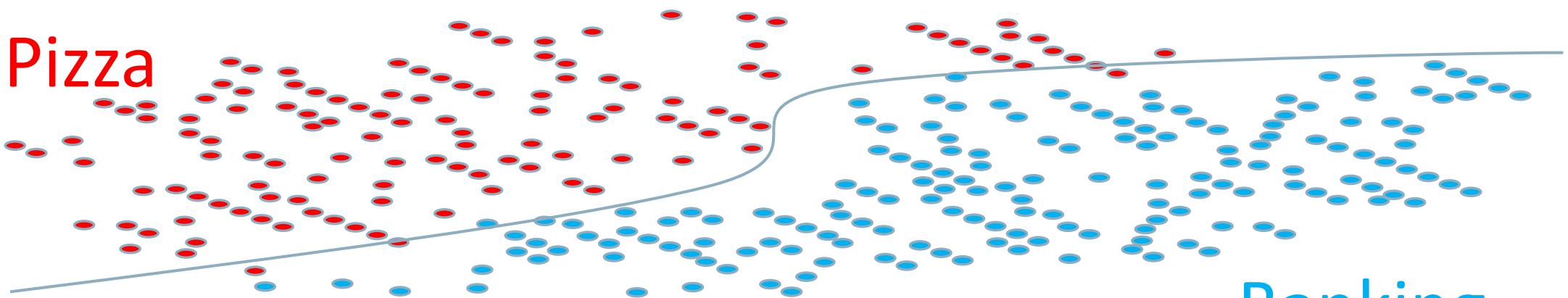


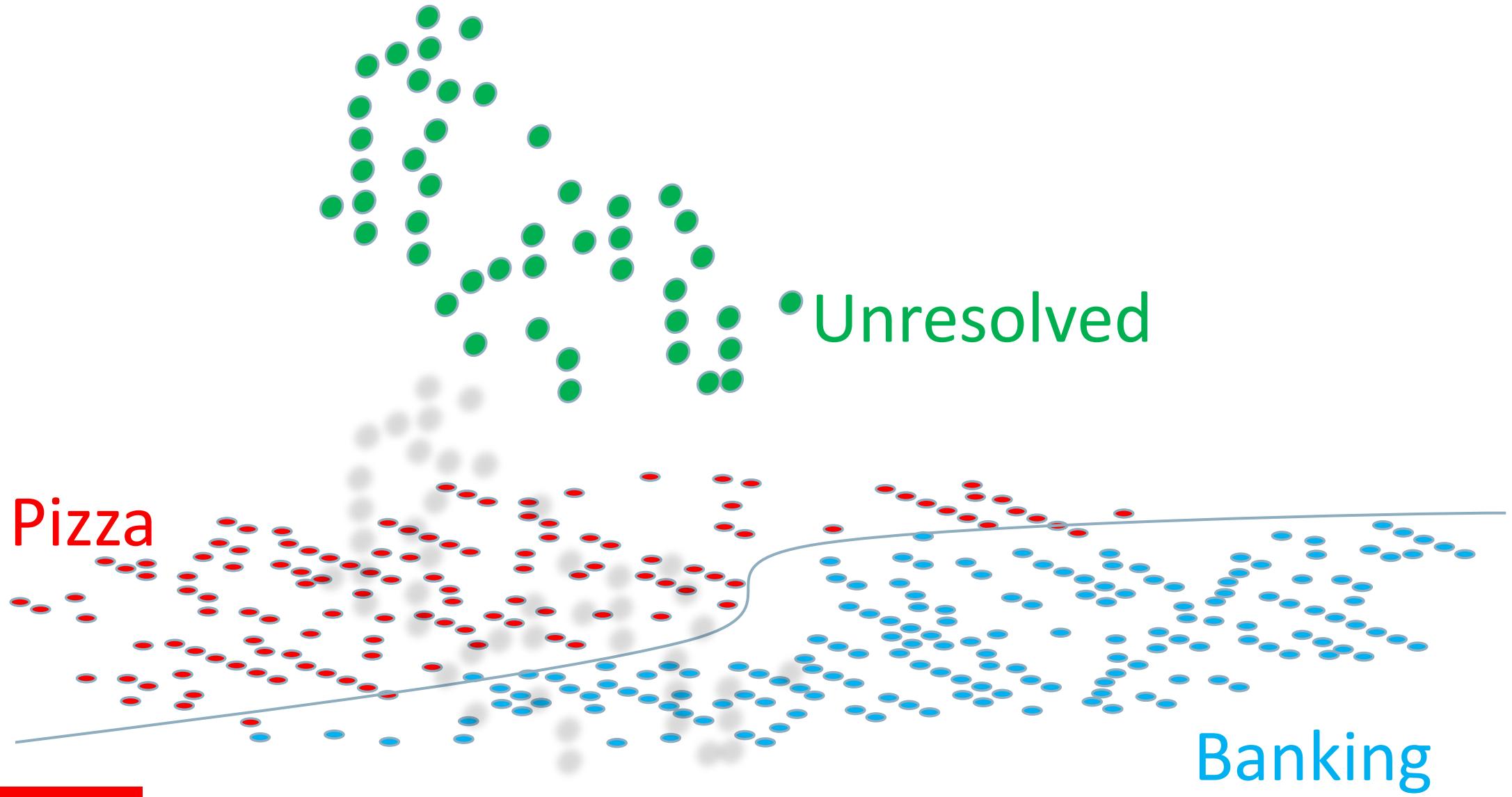
If your bot is only trained to know about **two possible intents**, then it will try to resolve any input to **one of those intents**

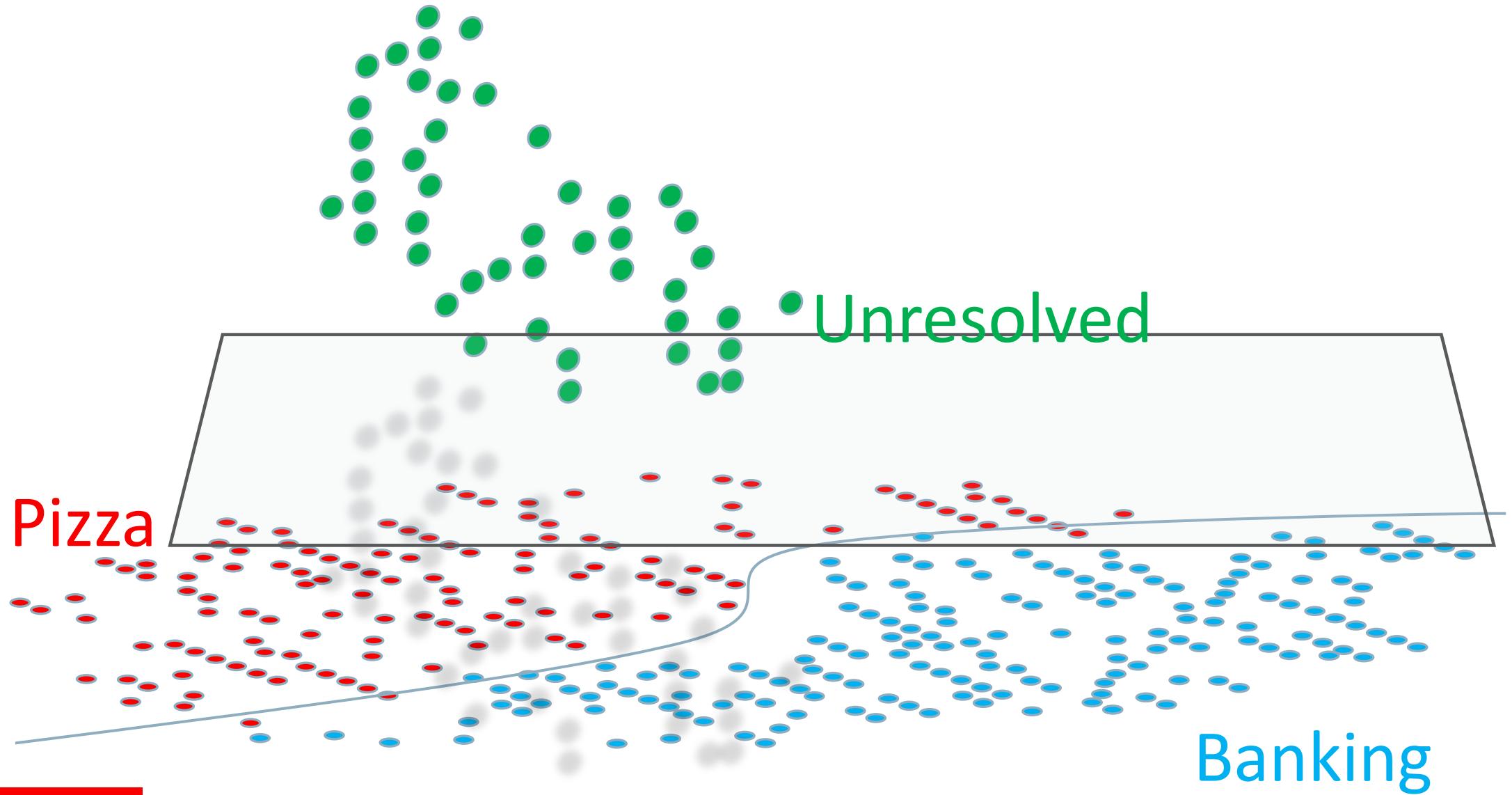


Pizza

Banking





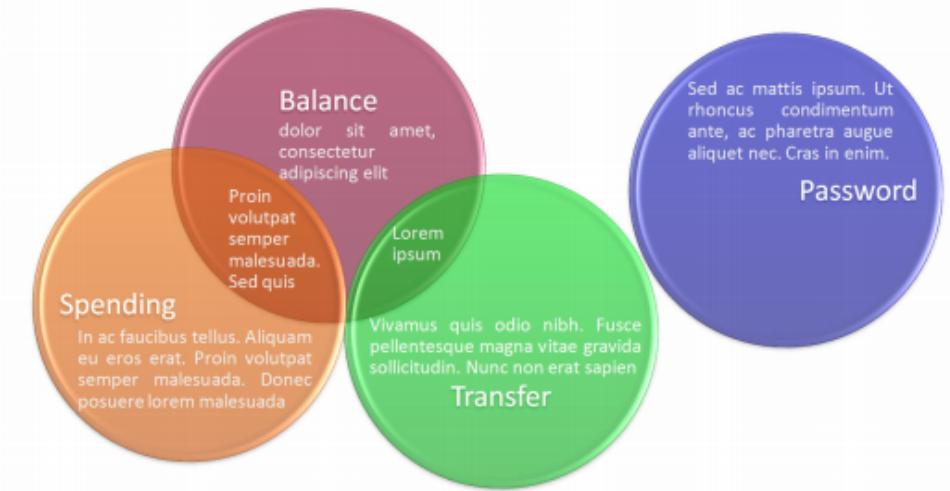


Topic agenda

- 1 ➤ Recap – the models in Oracle Digital Assistant
- 2 ➤ General guidelines for training your skills
- 3 ➤ Specific tips for training your skills
- 4 ➤ Training the digital assistant
- 5 ➤ Quality reports

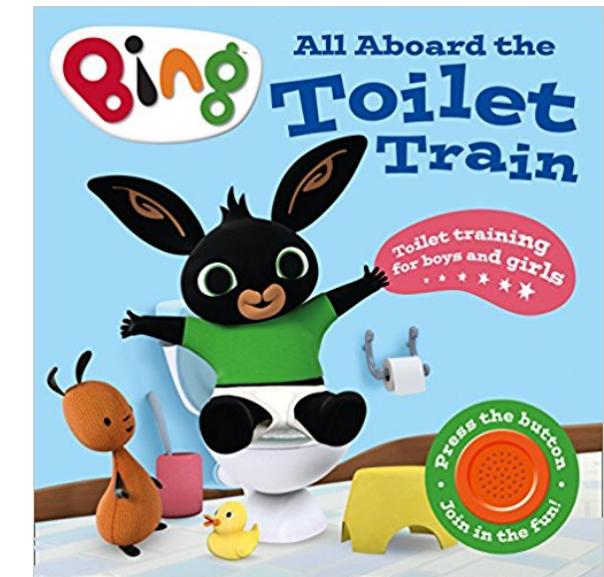
When synthesizing utterances what makes an intent unique

- The goal is to disambiguate intents
- What are key actors, actions and objects
 - “who” does “what” to “who”
 - Design utterances around these
- Consider combining intents if too similar
 - E.g. home, travel, car insurance
- Extra words can be used to weight utterances
 - But always retest when you do this



When synthesizing utterances be careful of lexical chrome

- Some words appear across all intents but add no meaning
 - “please”, “thank you”
- Ensure you don’t unintentionally skew results
 - E.g. if one intent consistently uses “please” that can skew results if the user says “please”
- Entity value in training utterances might inadvertently skew intent resolution
 - BOOK_ORDER_INTENT “Please order All Aboard the Toilet Train”
 - Incorrectly resolves “Does the train have a toilet”



When synthesizing utterances consider spelling & grammar

- Don't have to cover every form a word (Trainer Ht)
 - Different forms of a word are reduced to a common root
 - "want", "wanted" / "run" "ran" "running" "runs"
- Sentence structure may aid intent resolution (Trainer Ht)
 - Therefore use grammatically correct sentence where possible
- However, your data should still reflect your users' language
 - "wanna", "gonna", "gotta", "dunno"

When you see commonality consider combining intents

- More commonality of terms between intent utterances, the more difficult to disambiguate intents
- Consider creating one common intent and use entities to understand the discriminating factors

Updating Policy to add new person

Updating Policy to protect no claims

Add wife to my policy

Alter my insurance to add son

Can I change my policy to include spouse

Add my wife to insurance

Alter my policy to add wife

Update policy to include son

Add no claims to my policy

Alter my insurance to add protection

Can I change my policy to protect NCB

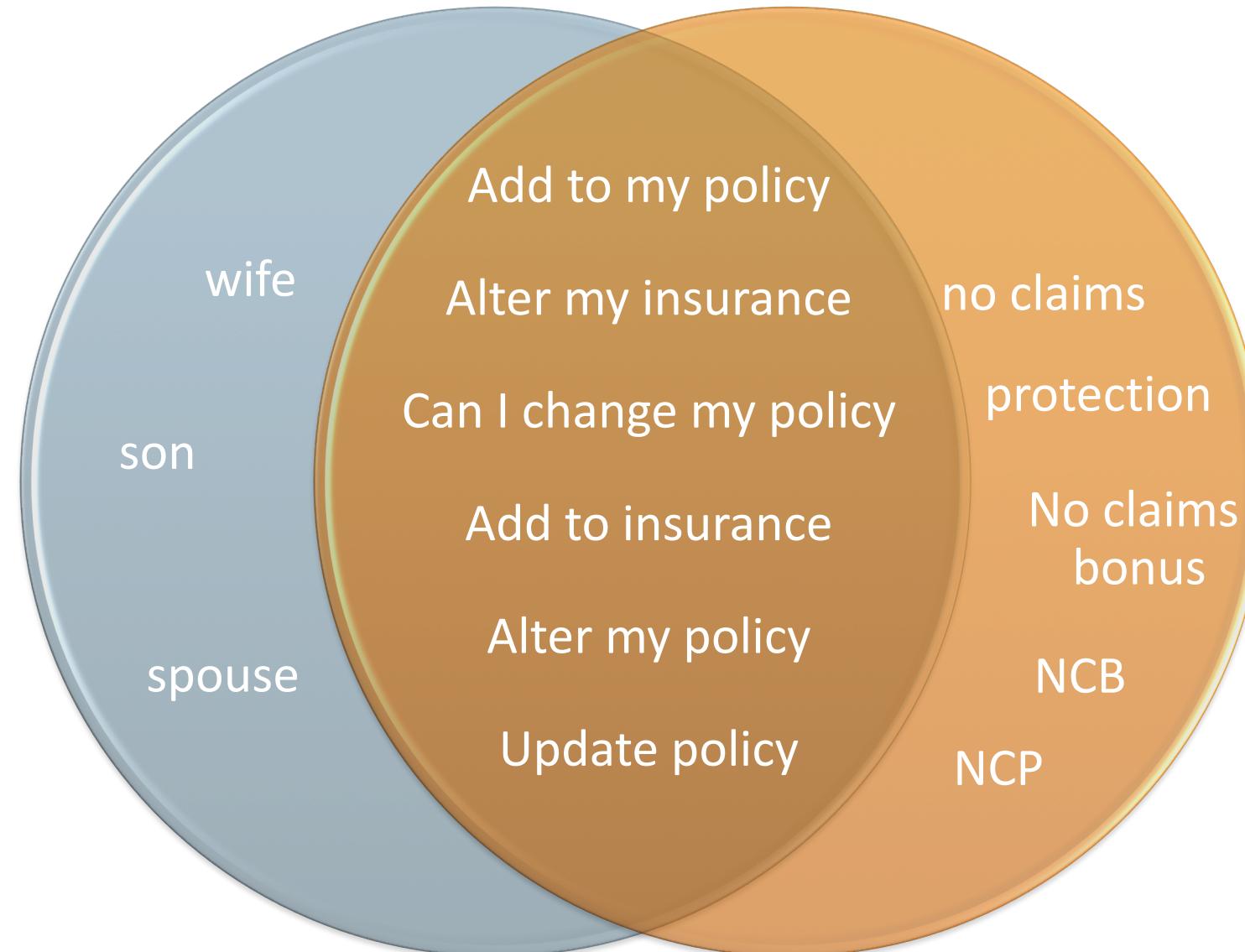
Add no claims protection to insurance

Alter my policy to add no claims bonus

Update policy to include NCP

Updating Policy to add new person

Updating Policy to protect no claims



Updating Policy

Person Entity

spouse

wife

son

Add to my policy

Alter my insurance

Can I change my policy

Add to insurance

Alter my policy

Update policy

Policy Claim Entity

No claims/

protection

No claims bonus

NCB/NCP

Tips for designing intents

- Create intents for the things you know the skill can't handle but is likely to be asked
 - Smalltalk, swearing, common business uses cases not handled by the skill
- A single use case might be implemented by multiple intents
 - Returns policy might be handled by 3 different intents (all execute the same flow)
 - Direct questions about company returns policy
 - Delivery being damaged
 - Clothes purchases wrong size
 - Potentially cleaner classification
 - Opportunity to handle each intent differently in the future

Train the unresolved intent

- Capture phrases which your bot should consider unresolved
 - Create unresolvedIntent intent
 - Use this to record any phrases you DON'T want to be resolved to your actual intents
- Analyze bot conversations and train the bot with any malicious/mismatched/rogue input
- Gives the intent resolution engine knowledge of what is NOT an intent

Topic agenda

- 1 ➤ Recap – the models in Oracle Digital Assistant
- 2 ➤ General guidelines for training your skills
- 3 ➤ Specific tips for training your skills
- 4 ➤ Training the digital assistant
- 5 ➤ Quality reports

Training the digital assistant

- All training guidelines apply to the skill
- There is limited NLP at the digital assistant level
 - Help
 - Exit
 - unresolvedIntent for disambiguating help and exit only.
 - No need to extensively train DA-level unresolvedIntent
- Digital assistant treats a skill unresolvedIntent in a special way
 - If no active skill, shows all available skills in the DA automatically
 - If a skill is active, the skill's unresolvedIntent handles the user input

Topic agenda

- 1 ➤ Recap – the models in Oracle Digital Assistant
- 2 ➤ General guidelines for training your skills
- 3 ➤ Specific tips for training your skills
- 4 ➤ Training the digital assistant
- 5 ➤ Quality reports

Introduction to quality reports

- Report that highlights possible intent resolution errors
 - Where utterances within intents are too similar
- Use a random 20% of the utterances to test the other 80%
- Compares all possible combination of pairs of intents in order to report
 - High quality: intents are distinct
 - Medium quality: some similar utterances
 - Low quality: the intent pairs aren't differentiated enough
- Indicates the number of utterances which may be problematic



[Int_Agnt_Exm_Authrty](#)
 \leftrightarrow [Int_Agnt_Min_Age](#)



[Int_Agnt_Shift_Agncy](#)
 \leftrightarrow [Int_Agnt_Shift_Cmpny](#)

2



[Int_Agnt_Min_Age](#)
 \leftrightarrow [Int_Agnt_Rtrmnt_Age](#)

4

Quality report on utterances

- Indicates utterances that are deemed to be similar
- Misclassified utterances (tests did not resolve to correct intent)

The screenshot shows the Oracle Quality Reports interface. The top navigation bar includes a back arrow, the title 'GR_QualityReports', and various navigation links like 'Instant Apps', 'Validate', 'Train', and a play button. On the left, there's a sidebar with icons for different report types. The main content area has tabs for 'Utterances', 'Suggestions', and 'History', with 'Utterances' selected. A 'Rerun Report' button and a 'Show All' toggle are at the top left. Below that is a 'Filter' search bar and a 'Sort By' dropdown set to 'Quality Ascending'. A 'Similar Utterances' section lists pairs of utterances with their expected and observed intents. A 'Misclassified Utterances' section lists individual utterances with their expected and observed intents and accuracy scores. The interface uses a light blue and white color scheme with red and yellow icons for error and warning status.

| Utterance | Expected Intent | Observed Intent | Accuracy |
|--|------------------------------|------------------------------|----------|
| When will i get monthly commission statement | Int_Agent_Periodicity_Paymnt | Int_Agent_Monthly_Statmnt | ! |
| Will i get my commission every month | Int_Agent_Periodicity_Paymnt | Int_Agent_Monthly_Statmnt | ! |
| When will i get monthly commission statement | Int_Agent_Monthly_Statmnt | Int_Agent_Periodicity_Paymnt | ! |

Quality report suggestions

- Reports any suggested changes you can make to your intents or utterances

The screenshot shows a user interface for quality reporting. At the top, there's a blue header bar with the title "GR_QualityReports" and several navigation icons: Instant Apps, Validate, Train, and a refresh symbol. Below the header is a toolbar with icons for Utterances, Suggestions, and History, along with a "Rerun Report" button and a "Filter" search bar. A sidebar on the left contains various icons for different report types. The main content area is titled "Suggestions" and displays a message: "There should be 5 or more utterances defined for the Empty Intent intent." Below this message, there's a section titled "Items with Suggestions" which lists one item: "Empty Intent". At the bottom of the page, it says "Page 1 of 1 (1 of 1 items)" with navigation arrows.

Quality report history

- View past user input ranked by win margin and confidence level
- Useful for
 - Narrow margins between intent classification
 - Intent resolution failures
 - Low confidence resolution

The screenshot shows the 'GR_QualityReports' application interface. On the left is a sidebar with icons for Home, Utterances, Suggestions, History, Instant Apps, Validate, Train, and Help. The main area has tabs for Utterances, Suggestions, and History, with History selected. A search bar at the top says 'Show me all customer messages Last 30 Days' with a date range '2/3/18 - 3/5/18'. Below it, a filter for 'Win Margin Is Less Than 10%' is shown. A green button labeled '+ Criteria' is visible. To the right, there are two sections: 'Customer Message' (listing five messages) and 'Intent Data' (listing two intents with their win margin and confidence levels). At the bottom, there's a 'General' section with timestamp and channel information.

| Customer Message | Intent | Win Margin | Confidence |
|--|--------------------------------|------------|------------|
| what do i do once I passed test | Int_Agent_After_Exam_Procedure | 3.9% | 18% |
| where can I sit an exam | Int_Agent_Passing_Marks | N/A | 14% |
| Is there a method for applying to be a specialist on the web | | | |
| can I add my wife | | | |
| how much to be an agent | | | |

Intent Data:

| Intent | Win Margin | Confidence |
|--------------------------------|------------|------------|
| Int_Agent_After_Exam_Procedure | 3.9% | 18% |
| Int_Agent_Passing_Marks | N/A | 14% |

General:

| Timestamp | Channel |
|-----------|---------|
| 3/2/18 | test |

Quality report history

- Show top intents that resolved with low confidence

Top Intent Confidence ▾ Is Less Than ▾ 50% ▾ ▾ ×

- Show where there may have been ambiguity between intents

Win Margin ▾ Is Less Than ▾ 20% ▾ ▾ ×

- Show me used input where the top intent was unresolved

Top Intent Name ▾ Is Equal To ▾ unresolvedIntent ×



Oracle Digital Assistant Hands-On

TBD

ORACLE®