NLP Course - Lesson 1

Group Exercise: Predicting the Next Word

Natural Language Processing

Today's Challenge: How Do We Predict the Next Word?

Your Mission

Work in groups of 3-4 students to figure out how YOU would predict the next word in a sentence.

No computers needed - just your brain!

Think about:

- How do YOU know what word comes next when you read?
- What patterns do you notice?
- Can you create rules or a method?

Your Training Text

The Dragon Story

Once upon a time, in a land filled with green meadows and tall mountains, there lived a friendly dragon. This friendly dragon loved to fly over the green meadows every morning. The villagers would watch the friendly dragon and wave. The dragon was not like other dragons; this dragon was kind and gentle. One day, a lost knight stumbled upon the land. The knight was afraid of the dragon, but the friendly dragon offered the knight a warm smile. The knight, seeing the kind dragon, was no longer afraid. The knight and the dragon became the best of friends, and they would often fly over the green meadows together.

Warm-Up: Complete These Sentences

Based on the story, what word would you predict comes next?

1. "The friendly _____"
2. "The knight and the _____"
3. "fly over the _____"
4. "The dragon was _____"
5. "in a land filled with _____"

Question: How did you decide? Write down your reasoning!

Group Exercise Part 1: Count the Patterns (15 minutes)

Task 1: Manual Counting

In your group, create a simple table counting what words follow these common words:

Word	Next Word	Count
the	dragon	?
the	friendly	?
the	knight	?
the	green	?
the	...	?
friendly	dragon	?
friendly	...	?
dragon	was	?
dragon	loved	?
dragon	...	?

Tip: Focus on the most frequent word pairs!

Group Exercise Part 2: Design Your Method (20 minutes)

Task 2: Create Your Prediction Algorithm

Now design a METHOD to predict the next word. Consider:

1. **Simple Approach:**
   - Look at the last word only?
   - What word most often follows it?
2. **Better Approach:**
   - Look at the last 2 words?
   - Look at the last 3 words?
3. **Your Creative Ideas:**
   - Consider word types (nouns, verbs)?
   - Consider sentence position?
   - Other patterns you noticed?

Write your method as simple steps (like a recipe)!

Group Exercise Part 3: Test Your Method (15 minutes)

Task 3: Make Predictions

Use your method to predict the next word:

1. "The friendly dragon loved to _____"
2. "The villagers would _____"
3. "The knight was _____"
4. "They would often _____"
5. "Once upon a _____"

Evaluate Your Method

- Does it give reasonable predictions?
- When does it work well?
- When does it fail?
- How could you improve it?

Think Deeper: The Challenges

Discussion Questions (10 minutes)

Discuss in your group:

1. **Memory Problem:**
   What if our story was 1000 times longer? How would you count all the patterns?
2. **New Words Problem:**
   What if you see: "The purple dragon..."
   You've never seen "purple" before. What now?
3. **Context Problem:**
   "The dragon was..." could be followed by many words.
   How do you choose the BEST one?
4. **Meaning Problem:**
   "friendly" and "kind" mean similar things.
   Can your method understand this?

Real-World Application

Where is Next-Word Prediction Used?

- **Your Phone:** Auto-complete when texting
- **Google:** Search suggestions
- **ChatGPT:** Generating entire conversations
- **Translation:** Converting between languages
- **Writing Assistants:** Helping write emails

The Big Insight

All these systems started with the SAME problem you just solved:
"Given some words, predict what comes next"

Share Your Solutions

Group Presentations (2 minutes each)

Each group presents:

1. Your prediction method (in 3-4 simple steps)
2. One example where it works well
3. One example where it fails
4. Your biggest challenge

Class Discussion

- Which methods were similar?
- Which were creative/different?
- What problems did everyone face?

Reflection: What You've Discovered

Key Insights

Through this exercise, you've discovered:

- **Patterns Matter:** Language has patterns we can count
- **Context Helps:** More context = better predictions
- **Frequency Works:** Common patterns are good predictors
- **It's Hard!:** Simple rules aren't enough for real language

What's Next?

In the coming lessons, we'll learn how computers solve these exact problems:

- N-gram models (counting patterns like you did!)
- Neural networks (learning patterns automatically)
- Word embeddings (understanding meaning)
- Transformers (what ChatGPT uses)

Homework: Observe Language Patterns

Before Next Class

1. **Observation Task:**
   Pay attention to your phone's autocomplete for one day.
   Write down 3 times it was right and 3 times it was wrong.
2. **Think About:**
   What patterns does it seem to use?
   Why does it fail when it fails?
3. **Bring to Next Class:**
   Your examples and observations

Remember

You just solved the fundamental problem of NLP!
Everything else builds on what you discovered today.

Instructor Notes

Exercise Goals

- Students discover n-gram concept naturally
- Experience the challenges of language modeling
- Build intuition before formal methods
- Encourage creative thinking

Expected Student Discoveries

- Most will create bigram/trigram-like methods
- Some might notice word categories
- All will struggle with unseen combinations
- Discussion reveals need for better methods

Time Management

- Total: 90 minutes
- Counting: 15 min
- Method Design: 20 min
- Testing: 15 min
- Discussion: 10 min
- Presentations: 20 min

Answer Key for Instructors

Most Frequent Next Words in the Text

"the" →	dragon (4), friendly (3), knight (3), green (2)
"friendly" →	dragon (4)
"dragon" →	loved (1), and (2), was (2), offered (1)
"green" →	meadows (3)
"knight" →	was (2), and (2), a (1)

Sample Predictions (Based on Frequency)

1. "The friendly **dragon**" (appears 4 times)
2. "fly over the **green**" or **land** (context dependent)
3. "The knight was **afraid**" or **no** (both appear)
4. "They would often **fly**" (from context)
5. "Once upon a **time**" (fixed phrase)