

Distant Supervision

- ▶ Manual labeling of the entire corpus is not possible
- ▶ Weak Supervision as a middle ground
- ▶ Use Heuristics to align a table of facts with the corpus
- ▶ Fuzzy training

Distant Supervision

Example

- ▶ Born - In database

Donald Knuth	Wisconsin
Srinivasa Ramanujan	Erode
Alan Turing	London

- ▶ Given Sentences

- ▶ Srinivasa Ramanujan was born in his maternal grandmother's home in Erode.
- ▶ Srinivasa Ramanujan was born in Erode, Tamilnadu, India, on 22nd December, 1887.
- ▶ Turing was born in Paddington, London, while his father was on leave from his position with the Indian Civil Service (ICS) at Chhatrapur, Bihar
- ▶ Alan Turing biopic The Imitation Game named as London film festival opener.

Distant Supervision

Example

- ▶ Born - In database

Donald Knuth	Wisconsin
Srinivasa Ramanujan	Erode
Alan Turing	London

- ▶ Given Sentences

- ▶ Srinivasa Ramanujan was born in his maternal grandmother's home in Erode. LABEL : BornIn
- ▶ Srinivasa Ramanujan was born in Erode, Tamilnadu, India, on 22nd December, 1887.
- ▶ Turing was born in Paddington, London, while his father was on leave from his position with the Indian Civil Service (ICS) at Chhatrapur, Bihar
- ▶ Alan Turing biopic The Imitation Game named as London film festival opener.

Distant Supervision

Example

- ▶ Born - In database

Donald Knuth	Wisconsin
Srinivasa Ramanujan	Erode
Alan Turing	London

- ▶ Given Sentences

- ▶ Srinivasa Ramanujan was born in his maternal grandmother's home in Erode.
- ▶ Srinivasa Ramanujan was born in Erode, Tamilnadu, India, on 22nd December, 1887.
- ▶ Turing was born in Paddington, London, while his father was on leave from his position with the Indian Civil Service (ICS) at Chhatrapur, Bihar
- ▶ Alan Turing biopic The Imitation Game named as London film festival opener.

Distant Supervision

Example

- ▶ Born - In database

Donald Knuth	Wisconsin
Srinivasa Ramanujan	Erode
Alan Turing	London

- ▶ Given Sentences

- ▶ Srinivasa Ramanujan was born in his maternal grandmother's home in Erode.
- ▶ Srinivasa Ramanujan was born in Erode, Tamilnadu, India, on 22nd December, 1887.
- ▶ Turing was born in Paddington, London, while his father was on leave from his position with the Indian Civil Service (ICS) at Chhatrapur, Bihar
- ▶ Alan Turing biopic The Imitation Game named as London film festival opener.

Distant Supervision

Example

- ▶ Born - In database

Donald Knuth	Wisconsin
Srinivasa Ramanujan	Erode
Alan Turing	London

- ▶ Given Sentences

- ▶ Srinivasa Ramanujan was born in his maternal grandmother's home in Erode.
- ▶ Srinivasa Ramanujan was born in Erode, Tamilnadu, India, on 22nd December, 1887.
- ▶ Turing was born in Paddington, London, while his father was on leave from his position with the Indian Civil Service (ICS) at Chhatrapur, Bihar
- ▶ Alan Turing biopic The Imitation Game named as London film festival opener.

Distant Supervision

- Distant supervision assumption, any sentence containing the entity pair will express the corresponding relation
- Can quickly label huge corpora
- Same entity pair can match different relations (Founded(Steve Jobs, Apple) or CEO(Steve Jobs, Apple))
- False positives, may lead to model learning wrong patterns for relations