What time is it? Temporal Analysis of Novels

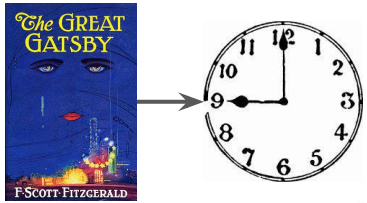
Allen Kim, Charuta Pethe, Steven Skiena
Data Science Lab

Stony Brook University
Computer Science

What we are trying to do

Predict **hour of day** (0-23) throughout a book

Important for understanding story flow and how events in a book are segmented



At ***nine o'clock, one morning*** late in July Gatsby's gorgeous car lurched up the rocky drive to my door and gave out a burst of melody from its three noted horn...

Prediction: Hour 9 (not 21 due to mention of morning)

By ***six o'clock*** Michaelis was worn out and grateful for the sound of a car stopping outside. It was one of the watchers of the night before who had promised to come back so he cooked ***breakfast*** for three which he and the other man ate together.

Prediction: Hour 6 (not 18 due to mention of breakfast)

Overview of Methods

1. Collect text with time references from dataset of books
2. Clean and label all time data extracted
   - Disambiguating AM vs PM - binary classifier
3. Train a model for time of day prediction for a window of text
   - 24-hour classifier
4. Use model to predict time over the entire book
5. Use dynamic programming to unify times throughout the book
   - Handling continuation of time

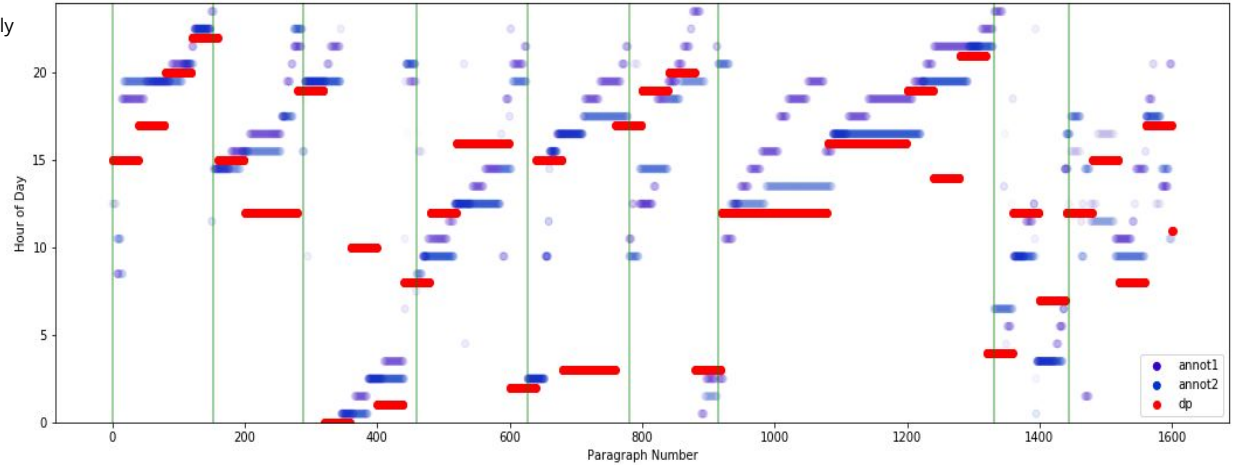
	Gutenberg		HathiTrust	
Hour	A.M	P.M	A.M	P.M
0	21,810	21,646	123,214	121,649
1	1,038	315	29,582	1,696
2	1,337	661	7,330	3,582
3	1,139	965	6,060	4,867
4	911	1032	5,133	5,150
5	645	720	4,053	3,693
6	754	562	4,316	3,125
7	531	495	3,421	2,602
8	698	596	3,950	3,401
9	657	587	3,795	3,468
10	745	661	3,952	3,897
11	410	460	2,055	2,737

Core Details and Results

1. Collect time phrases using SUTime, a time parser from Stanford CoreNLP on two book datasets (Gutenberg, HathiTrust)
2. Implement three different models to predict times - Naive Bayes, LSTM, BERT
3. Models are trained for AM/PM classification as well as 24 hour classification
4. Times are predicted over *n* windows and we run DP to unify them into *k* segments

$$\max_{i \in [1,n-k]} \left(f(n-i, k-1) + \max_{h \in [0,23]} \sum_{j=n-i}^{n} p_h[j] \right)$$

Flow of Time in *The Great Gatsby*



hour	agreement	hour	agreement
0	0.79	6	0.76
1	0.69	7	0.76
2	0.79	8	0.80
3	0.82	9	0.69
4	0.75	10	0.71
5	0.80	11	0.77
Mean Agreement = 0.761			

We had human annotators compare with our AM/PM predictions and show the agreement (above). We also show the error in hours (below) for our two datasets. BERT performs best overall.

hour	NB	LSTM	BERT
Gutenberg	4.69	4.72	**4.09**
HathiTrust	4.38	3.36	**2.28**

The graph on the left shows both human-annotated (blue) and our generated predicted times with DP (red) throughout *The Great Gatsby*. The errors in hours are shown below.

Inter-annotator Error (annot1, annot2 in blue)	1.85 hours
Model Error (dp in red)	2.62 hours