

Materials supplied by Microsoft Corporation may be used for internal review, analysis or research only. Any editing, reproduction, publication, rebroadcast, public showing, internet or public display is forbidden and may violate copyright law.



Aspect and Sentiment Unification Model

ACM Web Search and Data Mining 2011

Yohan Jo & Alice Oh

alice.oh@kaist.edu

Users & Information Lab

KAIST

December 2010

Our Research

- KAIST: major research and undergrad/graduate education in Korea
 - KAIST CS has 49 full-time tenure-track faculty
- Research at Users & Information Lab
 - Topic modeling: LDA, HDP and their variants
 - Sentiment analysis of reviews, Twitter, and other user-generated contents
- We welcome collaborations and discussions: email alice.oh@kaist.edu

Our Research

- KAIST: major research and undergrad/graduate education in Korea
 - KAIST CS has 49 full-time tenure-track faculty
- Research at Users & Information Lab
 - Topic modeling: LDA, HDP and their variants
 - Sentiment analysis of reviews, Twitter, and other user-generated contents
- We welcome collaborations and discussions: email alice.oh@kaist.edu

Problem: Unstructured reviews

363 of 381 people found the following review helpful:

★★★★★ My Powershot S90 Has Arrived, and It's Fantastic!, August 29, 2009

By Ray "A Reader"  - See all my reviews

TOP 500 REVIEWER

Amazon Verified Purchase [\(What's this?\)](#)

This review is from: Canon PowerShot S90 10MP Digital Camera with 3.8x Wide Angle Optical Image Stabilized Zoom and 3-inch LCD (Electronics)

Early Impressions

I was delighted to have finally received my Powershot S90, and without further adieu, let me say that this is one heck of a camera. It's not perfect -- you can take truly terrible pictures with it just like you can take terrible pictures with a D700 -- but when used properly, the camera turns out remarkable shots that make us find it hard to accept the images are coming from a camera that fits in your pants pocket.

What struck me first upon using it? First, it's size. This thing is small, and it's light, too. It's a bit smaller than my Panasonic TZ3 and TZ5, and it's lighter, too. (The camera uses a front and back metal construction with plastic on the top and bottom, but the Panasonic's, while also using metal, use a thicker gauge steel which adds a feel of sturdiness but also adds some weight, as well). The camera also has a high-quality feel to it. The buttons click and depress well (although the rear wheel is a bit too easy to turn, in my opinion). It has a rounded shape, so it feels comfortable in the hands, and when you stick it in your pocket, it will slide right in and out without snagging. The screen on this thing is simply gorgeous: why can't every camera have a screen like this? It's large, bright, and pretty high in resolution (461,000 pixels). You can't help but admire the camera's design once you get looking at it and using it.

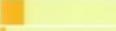
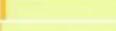
Next, the camera seems to perform well in terms of speed and overall operational use. The screen has the typical lag when taking shots, but you can adjust this somewhat in the menu system to speed things up, and quite frankly, every small camera I've ever owned exhibits this behavior. It is easy to use most of the camera's functions, and you may have heard about the programmable control ring around the lens on the front of the camera. It's operation is easy, solid (the ring "clicks" with detents at different positions), and, to boot, there is the standard programmable "S" button that the Powershots "S" cameras have typically had.

But of course, I'm interested in high ISO operation, and so I immediately took it into the livingroom where it was quite dark, and just started shooting. I was quite surprised at the results. You are not going to necessarily submit these to win any contests, but for the most part, the camera took nice shots even in that bad shooting environment, and the vast majority of the photos came out quite well (I will post a few with this review). The camera is the first (along with the Powershot G11) to deploy Sony's new ICX685CQZ sensor, a 9.31mm diagonal sensor with high performance specifications. With a little post processing, many of them look quite good. As the ISO crept into the very high ranges (800 and above) some sensor noise became apparent, but this is certainly the best low-light performance I've seen in a non-DSLR so far. (See my explanations, below, to see why this is possible). Surprisingly, some shots as high as 1600 ISO seemed to be acceptable as long as you are not a "pixel peeper." I was quite surprised when I discovered that a few of the shots had been taken at this high ISO 1600 level -- I've never had this experience before with a point and shoot camera.

Outdoor operation is fantastic. My outdoor shots for the most part have come out very well, with rich color, great detail, and little sensor noise. Like most Canon portables, these images seem to respond well to post-processing (you can sharpen them quite easily, and Canon now uses a standard meta-data tagging format that is readable by virtually all photo editor programs.) I even turned the EV down -2/3 while outside, and the sensitivity of the camera is so good that, even with this reduction in EV, my shots came out sharp and clear. Again, I will post a few shots with this review.

Average Customer Rating

★★★★★ (233 customer reviews)

<u>5 star:</u>	 (150)
<u>4 star:</u>	 (49)
<u>3 star:</u>	 (12)
<u>2 star:</u>	 (13)
<u>1 star:</u>	 (9)

Customer Ratings

[Close](#)

Features	 (270)
Picture quality	 (269)
Portability	 (269)
Ease of use	 (268)
Battery life	 (260)
Low light pictures	 (257)

› See what customers say about these attributes

These aspects and aspect-specific sentiments are available on some Web sites for some of the products.

Customer Reviews

Canon PowerShot SD4000IS 10 MP CMOS Digital Camera with 3.8x Opti

Average Customer Rating

★★★★☆ (57 customer reviews)

<u>5 star:</u>	 (26)
<u>4 star:</u>	 (15)
<u>3 star:</u>	 (4)
<u>2 star:</u>	 (5)
<u>1 star:</u>	 (7)

► Be the first to [suggest an attribute for this item.](#)

[Create your own review](#)

Can we automatically find and analyze the relevant attributes and the aspect-specific sentiments?

e finally received my Powershot S90, and without further adieu, let me say that this is one heck of a camera. I can take truly terrible pictures with it just like you can take terrible pictures with a D700 -- but when used it turns out remarkable shots that make us find it hard to accept the images are coming from a camera that fits

pon using it? First, it's size. This thing is small, and it's light, too. It's a bit smaller than my Panasonic TZ3 and D90. (The camera uses a front and back metal construction with plastic on the top and bottom, but the body using metal, use a thicker gauge steel which adds a feel of sturdiness but also adds some weight, as well). It has a high-quality feel to it. The buttons click and depress well (although the rear wheel is a bit too easy to turn, in a rounded shape, so it feels comfortable in the hands, and when you stick it in your pocket, it will slide right in). The screen on this thing is simply gorgeous: why can't every camera have a screen like this? It's large, in resolution (461,000 pixels). You can't help but admire the camera's design once you get looking at it and

ns to perform well in terms of speed and overall operational use. The screen has the typical lag when taking photos, but this somewhat in the menu system to speed things up, and quite frankly, every small camera I've ever used has this behavior. It is easy to use most of the camera's functions, and you may have heard about the programmable ring on the lens on the front of the camera. Its operation is easy, solid (the ring "clicks" with detents at different positions), there is the standard programmable "S" button that the Powershots "S" cameras have typically had.

I rested in high ISO operation, and so I immediately took it into the livingroom where it was quite dark, and I was quite surprised at the results. You are not going to necessarily submit these to win any contests, but for a camera took nice shots even in that bad shooting environment, and the vast majority of the photos came out quite well (I'll show a few with this review). The camera is the first (along with the Powershot G11) to deploy Sony's new 1/1.7" 9.31mm diagonal sensor with high performance specifications. With a little post processing, many of them at higher ISO crept into the very high ranges (800 and above) some sensor noise became apparent, but this is excellent high performance I've seen in a non-DSLR so far. (See my explanations, below, to see why this is possible). The shots as high as 1600 ISO seemed to be acceptable as long as you are not a "pixel peeper." I was quite impressed that a few of the shots had been taken at this high ISO 1600 level -- I've never had this experience with a point-and-shoot camera.

I am quite fantastic. My outdoor shots for the most part have come out very well, with rich color, great detail, and little noise. Compared to Canon portables, these images seem to respond well to post-processing (you can sharpen them quite easily, and they retain their sharpness). The camera uses a standard meta-data tagging format that is readable by virtually all photo editor programs.) I even turned the flash off outside, and the sensitivity of the camera is so good that, even with this reduction in EV, my shots came out quite well. I will post a few shots with this review.

e finally received my Powershot S90, and without further adieu, let me say that this is one heck of a camera. can take truly terrible pictures with it just like you can take terrible pictures with a D700 -- but when used turns out remarkable shots that make us find it hard to accept the images are coming from a camera that fits

pon using it? First, it's size. This thing is small, and it's light, too. It's a bit smaller than my Panasonic TZ3 and bo. (The camera uses a front and back metal construction with plastic on the top and bottom, but the o using metal, use a thicker gauge steel which adds a feel of sturdiness but also adds some weight, as well). a high-quality feel to it. The buttons click and depress well (although the rear wheel is a bit too easy to turn, in rounded shape, so it feels comfortable in the hands, and when you stick it in your pocket, it will slide right in ing. The screen on this thing is simply gorgeous why can't every camera have a screen like this? It's large, in resolution (461,000 pixels). You can't help but admire the camera's design once you get looking at it and

ns to perform well in terms of speed and overall operational use. The screen has the typical lag when taking ust this somewhat in the menu system to speed things up, and quite frankly, every small camera I've ever havior. It is easy to use most of the camera's functions, and you may have heard about the programmable e lens on the front of the camera. Its operation is easy, solid (the ring "clicks" with detents at different t, there is the standard programmable "S" button that the Powershots "S" cameras have typically had.

rested in high ISO operation, and so I immediately took it into the livingroom where it was quite dark, and I was quite surprised at the results. You are not going to necessarily submit these to win any contests, but for hera took nice shots even in that bad shooting environment, and the vast majority of the photos came out few with this review). The camera is the first (along with the Powershot G11) to deploy Sony's new 9.31mm diagonal sensor with high performance specifications. With a little post processing, many of them e ISO crept into the very high ranges (800 and above) some sensor noise became apparent, but this is light performance I've seen in a non-DSLR so far. (See my explanations, below, to see why this is possible). pts as high as 1600 ISO seemed to be acceptable as long as you are not a "pixel peeper." I was quite overed that a few of the shots had been taken at this high ISO 1600 level -- I've never had this experience d shoot camera.

fantastic. My outdoor shots for the most part have come out very well, with rich color, great detail, and little st Canon portables, these images seem to respond well to post-processing (you can sharpen them quite easily, a standard meta-data tagging format that is readable by virtually all photo editor programs.) I even turned the tside, and the sensitivity of the camera is so good that, even with this reduction in EV, my shots came out , I will post a few shots with this review.

Overview of Talk

- Introduction to Topic Models
- LDA: Latent Dirichlet Allocation
- Aspect and sentiment in review data
- ASUM: Aspect and Sentiment Unification Model
- Experiments and results
 - Review data
 - Twitter data

Information overload



As more information becomes available, it becomes more difficult to find and discover what we need.

We need new tools to help us organize, search, and understand these vast amounts of information.

Topic modeling



Topic modeling provides methods for automatically organizing, understanding, searching, and summarizing large electronic archives.

- ① Discover the hidden themes that pervade the collection.
- ② Annotate the documents according to those themes.
- ③ Use annotations to organize, summarize, and search the texts.

Latent Dirichlet Allocation

Blei, Ng, and Jordan, JMLR 2003

1. Basic Assumption
2. Generative Process
3. Inference
4. Graphical Representation

Economic Slowdown Catches Up With Nascar

By KEN BELSON and DAVE CALDWELL

After years of jam-packed races, sky-high television ratings and record merchandise sales, Nascar has seen attendance at nearly every track slip this year as recession-weary fans continue to cut costs.

The Behler family could see that firsthand while sitting atop their old school bus in the infield at Pocono Raceway for last week's race in Long Pond, Pa. From that perch, they saw empty patches of grass with untrampled dandelions that in years past were covered by other spectators' cars, campers and trailers.

Fans like the Behlers who are showing up to races are spending less, too.

"Everybody's still coming, but no one's spending," said Susan Behler, who arrived last Sunday, race day, instead of Friday night to save money. "Three years ago, I used to spend \$200 or \$300 every time I came here. Now, it's a question: do I need it?"

Other sports leagues have been hurt the past two years. But Nascar — with its heavier reliance on working-class fans, low fuel prices and the beleaguered auto industry — has suffered disproportionately, racing industry executives say. Ratings on television, sales of licensed goods and sponsorships, the lifeblood of the sport, are also suffering. Several racing teams have merged in the last three years.

Nascar compounded matters, the executives say, by changing its rules in ways that made the racing safer but stripped the sport of some of the spontaneity that made it compelling. Under pressure, Nascar has reversed some of those moves, helping to rejuvenate competition on the track this season. With the economy on the mend, Nascar and its teams, sponsors, track owners and broadcasters seem confident that the worst is over.

nascar, races, track, raceway, race, cars, fuel, auto, racing

economic, slowdown, sales, recession, costs, spending, save

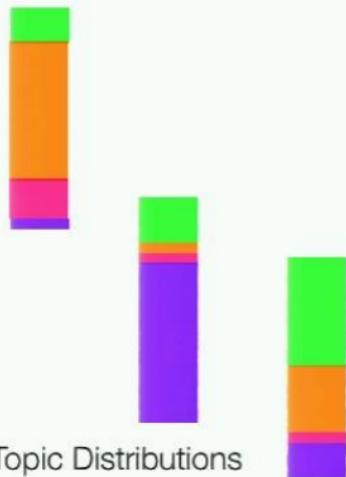
fans, spectators, sports, leagues, teams, competition

nascar, races, track, raceway, race, cars, fuel, auto, racing

economic, slowdown, sales, recession, costs, spending, save

fans, spectators, sports, leagues, teams, competition

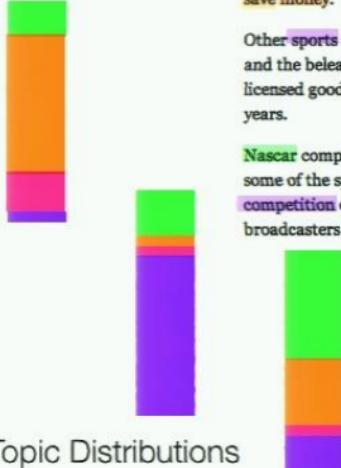
Topics: multinomial over words



Topic Distributions

nascar, races, track, raceway, race, cars, fuel, auto, racing
economic, slowdown, sales, recession, costs, spending, save
fans, spectators, sports, leagues, teams, competition

Topics: multinomial over words



August 8, 2010

<http://www.nytimes.com/2010/08/09/sports/autoracing/09nascar.html?>

Economic Slowdown Catches Up With Nascar

By KEN BELSON and DAVE CALDWELL

After years of jam-packed races, sky-high television ratings and record merchandise sales, Nascar has seen attendance at nearly every track slip this year as recession-weary fans continue to cut costs.

The Behler family could see that firsthand while sitting atop their old school bus in the infield at Pocono Raceway for last week's race in Long Pond, Pa. From that perch, they saw empty patches of grass with untrampled dandelions that in years past were covered by other spectators' cars, campers and trailers.

Fans like the Behlers who are showing up to races are spending less, too.

"Everybody's still coming, but no one's spending," said Susan Behler, who arrived last Sunday, race day, instead of Friday night to save money. "Three years ago, I used to spend \$200 or \$300 every time I came here. Now, it's a question: do I need it?"

Other sports leagues have been hurt the past two years. But Nascar — with its heavier reliance on working-class fans, low fuel prices and the beleaguered auto industry — has suffered disproportionately, racing industry executives say. Ratings on television, sales of licensed goods and sponsorships, the lifeblood of the sport, are also suffering. Several racing teams have merged in the last three years.

Nascar compounded matters, the executives say, by changing its rules in ways that made the racing safer but stripped the sport of some of the spontaneity that made it compelling. Under pressure, Nascar has reversed some of those moves, helping to rejuvenate competition on the track this season. With the economy on the mend, Nascar and its teams, sponsors, track owners and broadcasters seem confident that the worst is over.

nascar, races, track, raceway, race, cars, fuel, auto, racing

economic, slowdown, sales, recession, costs, spending, save

fans, spectators, sports, leagues, teams, competition

Topic Distributions

Topics: multinomial over words

Putting Ruth and Red Sox in the Rearview Mirror

By THOMAS KAPLAN

Derek Jeter has constructed a 16-year major league career out of clutch hits. So it was only fitting that he would provide several examples on the night when he surpassed [Ralph Ruth](#) with his 9,85th hit

August 8, 2010

<http://www.nytimes.com/2010/08/09/sports/autoracing/09nascar.html?>

It was also fitting that the Red Sox saw their

now find themselves By [KEN BELSON](#) and [DAVE CALDWELL](#)

The Yankees previously had been important. After years of jam-packed races, sky-high television ratings and record merchandise sales, Nascar has seen attendance at nearly every track slip this year as recession-weary fans continue to cut costs.

The Behler family could see that firsthand while sitting atop their old school bus in the infield at Pocono Raceway for last week's race in Long Pond, Pa. From that perch, they saw empty patches of grass with untrampled dandelions that in years past were covered by other spectators' cars, campers and trailers.

"You can't dream about fans like the Behlers who are showing up to races are spending less, too.

"Everybody's still coming, but no one's spending," said Susan Behler, who arrived last Sunday, race day, instead of Friday night to save money. "Three years ago, I used to spend \$200 or \$300 every time I came here. Now, it's a question: do I need it?"

The Red Sox hope them. They dream Now the best they six and a third inni Other sports leagues have been hurt the past two years. But Nascar — with its heavier reliance on working-class fans, low fuel prices and the beleaguered auto industry — has suffered disproportionately, racing industry executives say. Ratings on television, sales of licensed goods and sponsorships, the lifeblood of the sport, are also suffering. Several racing teams have merged in the last three years.

Nascar compounded matters, the executives say, by changing its rules in ways that made the racing safer but stripped the sport of some of the spontaneity that made it compelling. Under pressure, Nascar has reversed some of those moves, helping to rejuvenate competition on the track this season. With the economy on the mend, Nascar and its teams, sponsors, track owners and broadcasters seem confident that the worst is over.



Topic Distributions



nascar, races, track, raceway, race, cars, fuel, auto, racing
economic, slowdown, sales, recession, costs, spending, save
fans, spectators, sports, leagues, teams, competition

Topics: multinomial over words

Revenue Rises, and Chrysler Narrows Its Loss in Quarter

By NICK BUNKLEY

DETROIT — Chrysler said on Monday that it narrowed its loss in the second quarter to \$172 million and recorded its second consecutive operating profit as sales and revenue increased.

The carmaker ear
loss, primarily at
quarter loss of \$2

Putting Ruth and Red Sox in the Rearview Mirror

By THOMAS KAPLAN

Derek Jeter has constructed a 16-year major league career out of clutch hits. So it was only fitting that he would provide several more examples on the night Revenue rose 8.2 when he surpassed *Ruth* with his 9,874th hit

Chrysler increase
\$1.2 billion owed
The second quar
Sergio Marchion
grow."

It was also fitting t
the Red Sox saw th
now find themselv
The Yankees previa
Chrysler said in i important.

financial goals for
Their scheduled st
Mr. Marchionne c
rivals because its
profile counterpar

A little more than
production. It rec
"You can't dream a

August 8, 2010

<http://www.nytimes.com/2010/08/09/sports/autoracing/09nascar.html?>

Economic Slowdown Catches Up With Nascar

By KEN BELSON and DAVE CALDWELL

After years of jam-packed races, sky-high television ratings and record merchandise sales, Nascar has seen attendance at nearly every track slip this year as recession-weary fans continue to cut costs.

The Behler family could see that firsthand while sitting atop their old school bus in the infield at Pocono Raceway for last week's race in Long Pond, Pa. From that perch, they saw empty patches of grass with untrampled dandelions that in years past were covered by other spectators' cars, campers and trailers.

Fans like the Behlers who are showing up to races are spending less, too.

"Everybody's still coming, but no one's spending," said Susan Behler, who arrived last Sunday, race day, instead of Friday night to save money. "Three years ago, I used to spend \$200 or \$300 every time I came here. Now, it's a question: do I need it?"

Other sports leagues have been hurt the past two years. But Nascar — with its heavier reliance on working-class fans, low fuel prices and the beleaguered auto industry — has suffered disproportionately, racing industry executives say. Ratings on television, sales of licensed goods and sponsorships, the lifeblood of the sport, are also suffering. Several racing teams have merged in the last three years.

Nascar compounded matters, the executives say, by changing its rules in ways that made the racing safer but stripped the sport of some of the spontaneity that made it compelling. Under pressure, Nascar has reversed some of those moves, helping to rejuvenate competition on the track this season. With the economy on the mend, Nascar and its teams, sponsors, track owners and broadcasters seem confident that the worst is over.



Topic Distributions

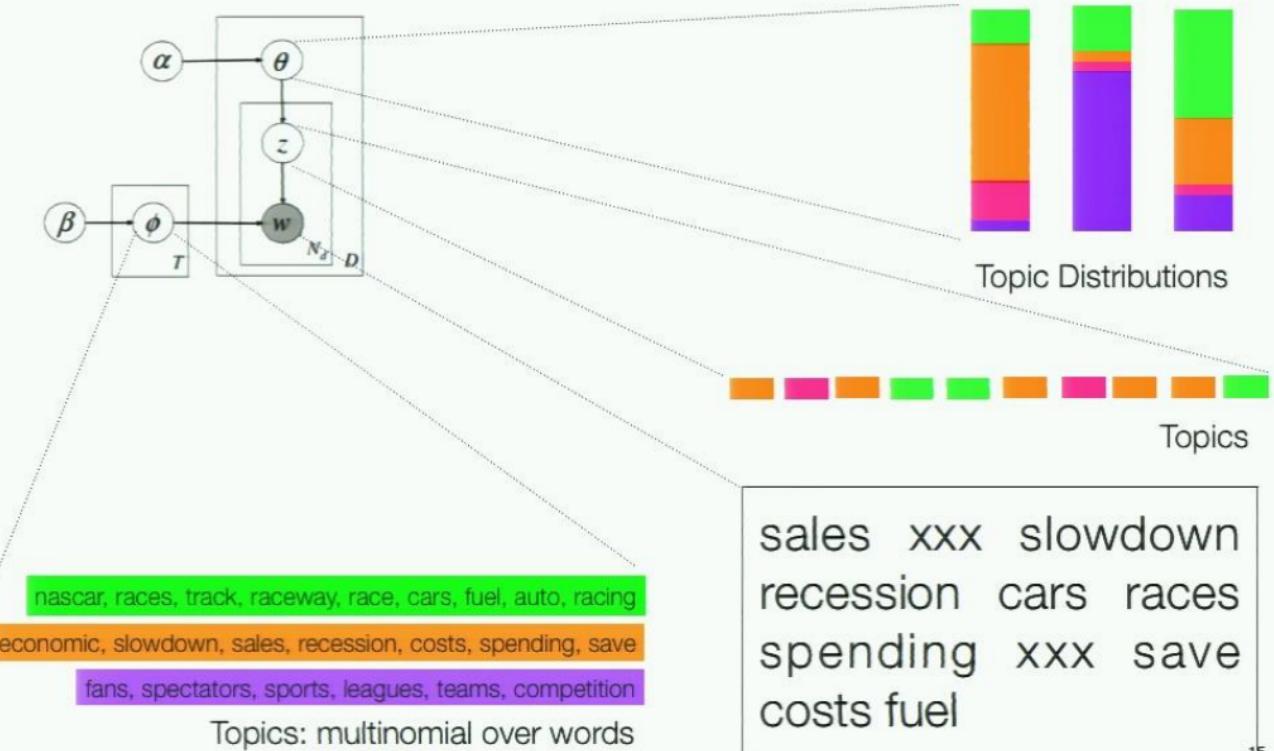
nascar, races, track, raceway, race, cars, fuel, auto, racing

economic, slowdown, sales, recession, costs, spending, save

fans, spectators, sports, leagues, teams, competition

Topics: multinomial over words

Graphical Representation of LDA



Input to LDA

Input to LDA

Revenue Rises, and Chrysler Narrows Its Loss in Quarter

By NICK BUNKLEY

DETROIT — Chrysler said on Monday that it narrowed its loss in the second quarter to \$172 million and recorded its second consecutive operating profit as sales and revenue increased.

The carmaker ear
loss, primarily at
quarter loss of \$2

Putting Ruth and Red Sox in the Rearview Mirror

By THOMAS KAPLAN

Derek Jeter has constructed a 16-year major league career out of clutch hits. So it was only fitting that he would provide several more examples on the night Revenue rose 8.2 when he surpassed Babe Ruth with his 5,746th hit

Chrysler increase
\$1.2 billion owed It was also fitting tha
the Red Sox saw the
“The second quar now find themselves
Sergio Marchionne
grow.”

August 8, 2010

The Yankees prevale
Chrysler said in t important.

financial goals for

Their scheduled start
Mr. Marchionne c
Dustin Moseley, did
rivals because its
profile counterpart, J

A little more than
production. It rec “You can’t dream any

Beckett’s return from
of boost they needed

The Red Sox hoped f
them. They dreamed

Now the best they ca
six and a third inning

<http://www.nytimes.com/2010/08/09/sports/autoracing/09nascar.html?>

Economic Slowdown Catches Up With Nascar

By KEN BELSON and DAVE CALDWELL

After years of jam-packed races, sky-high television ratings and record merchandise sales, Nascar has seen attendance at nearly every track slip this year as recession-weary fans continue to cut costs.

The Behler family could see that firsthand while sitting atop their old school bus in the infield at Pocono Raceway for last week's race in Long Pond, Pa. From that perch, they saw empty patches of grass with untrampled dandelions that in years past were covered by other spectators' cars, campers and trailers.

Fans like the Behlers who are showing up to races are spending less, too.

“Everybody’s still coming, but no one’s spending,” said Susan Behler, who arrived last Sunday, race day, instead of Friday night to save money. “Three years ago, I used to spend \$200 or \$300 every time I came here. Now, it’s a question: do I need it?”

Other sports leagues have been hurt the past two years. But Nascar — with its heavier reliance on working-class fans, low fuel prices and the beleaguered auto industry — has suffered disproportionately, racing industry executives say. Ratings on television, sales of licensed goods and sponsorships, the lifeblood of the sport, are also suffering. Several racing teams have merged in the last three years.

Nascar compounded matters, the executives say, by changing its rules in ways that made the racing safer but stripped the sport of some of the spontaneity that made it compelling. Under pressure, Nascar has reversed some of those moves, helping to rejuvenate competition on the track this season. With the economy on the mend, Nascar and its teams, sponsors, track owners and broadcasters seem confident that the worst is over.

Topics Discovered by LDA

nascar	0.12	spending	0.09	sports	0.12
races	0.10	economic	0.07	team	0.11
cars	0.10	recession	0.06	game	0.10
racing	0.09	save	0.05	player	0.10
track	0.08	money	0.05	athlete	0.09
speed	0.06	cut	0.04	win	0.07
...
money	0.002	speed	0.003	nascar	0.001

Topics: multinomial over vocabulary

Topic Distributions of Documents in the Corpus

Revenue Rises, and Chrysler Narrows Its Loss in Quarter

By NICK BUNKLEY

DETROIT — Chrysler said on Monday that it narrowed its loss in the second quarter to \$172 million and recorded its second consecutive operating profit as sales and revenue increased.

The carmaker earned \$18 million in operating income, which excludes one-time charges, an improvement of \$40 million over the first quarter. The overall loss, primarily at its

Putting Ruth and Red Sox in the Rearview Mirror

quarter loss of \$2.

By THOMAS KAPLAN

Derek Jeter has constructed a 16-year major league career out of clutch hits. So it was only fitting that he would provide several more examples on the night Revenue rose 8.2% when he surpassed Babe Ruth with his 2,876th hit.

Chrysler increased It was also fitting that Jeter's milestone hit came against the Boston Red Sox, one of Ruth's former teams. As Jeter saw his name rise on the career hits list, \$1.2 billion owned by the Red Sox saw their

now find themselves

Economic Slowdown Catches Up With Nascar

"The second quart

Sergio Marchionne The Yankees prevailed

grow."

By KEN BELSON and DAVE CALDWELL

After years of jam-packed races, sky-high television ratings and record merchandise sales, Nascar has seen attendance at nearly

Mr. Marchionne c

rivals because its

"You can't dream any

Pans like the Behlers who are showing up to races are spending less, too.

A little more than

production. It recr

Beckett's return from

of boost they needed

them. They dreamed

Now the best they ca

six and a third inning

years.

"Everybody's still coming, but no one's spending," said Susan Behler, who arrived last Sunday, race day, instead of Friday night to save money. "Three years ago, I used to spend \$200 or \$300 every time I came here. Now, it's a question: do I need it?"

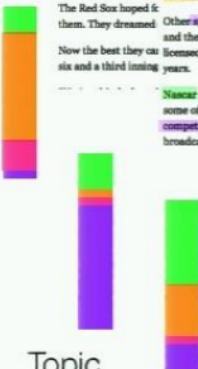
Other sports leagues have been hurt the past two years. But Nascar — with its heavier reliance on working-class fans, low fuel prices

and the beleaguered auto industry — has suffered disproportionately, racing industry executives say. Ratings on television, sales of

licensed goods and sponsorships, the lifeblood of the sport, are also suffering. Several racing teams have merged in the last three

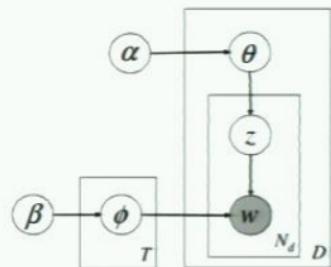
years.

Nascar compounded matters, the executives say, by changing its rules in ways that made the racing safer but stripped the sport of some of the spontaneity that made it compelling. Under pressure, Nascar has reversed some of those moves, helping to rejuvenate competition on the track this season. With the economy on the mend, Nascar and its teams, sponsors, track owners and broadcasters seem confident that the worst is over.

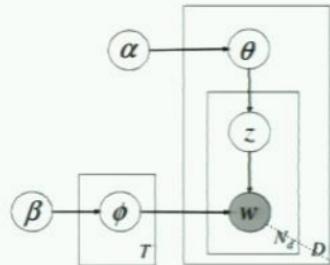


Topic distributions for each document in the corpus

Graphical View



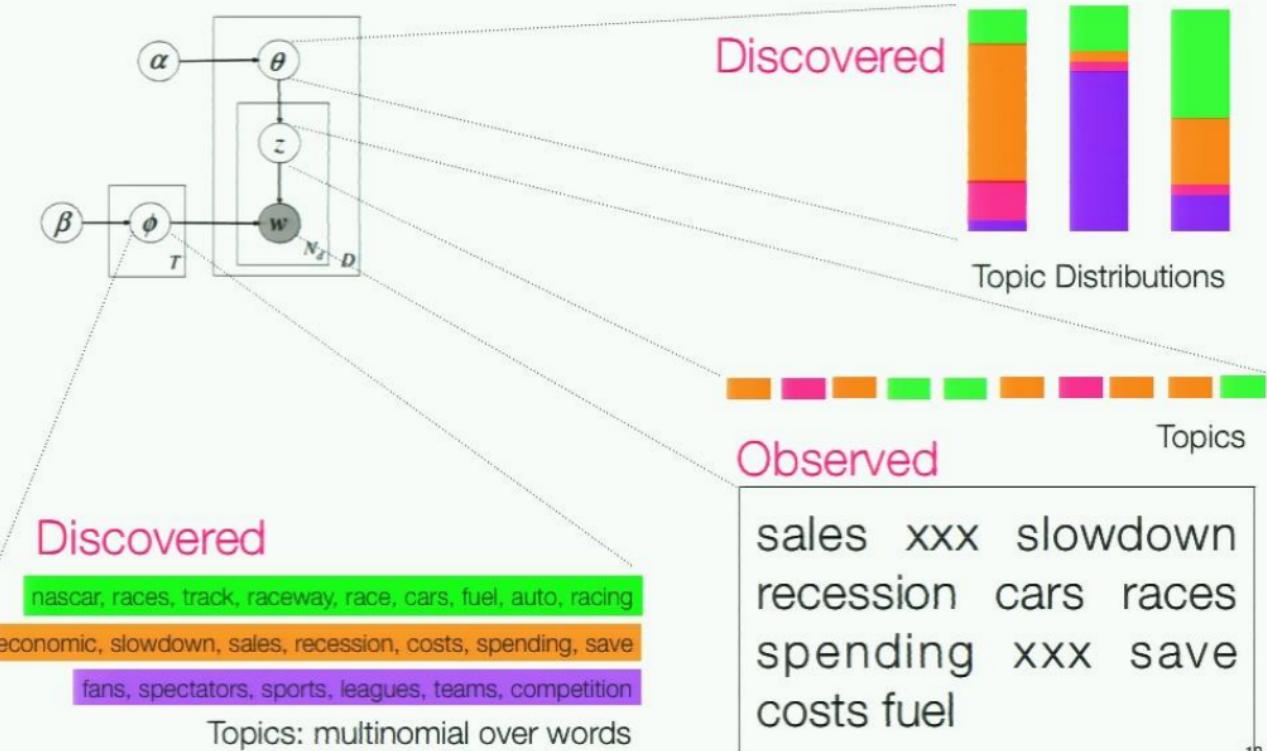
Graphical View



Observed

sales xxx slowdown
recession cars races
spending xxx save
costs fuel

Graphical View



ASUM: Aspect Sentiment Unification Model

to uncover the intertwined semantic structure of aspects and sentiments in reviews

Yohan Jo and Alice Oh
WSDM 2011

Problem

363 of 381 people found the following review helpful:

★★★★★ My Powershot S90 Has Arrived, and It's Fantastic!, August 29, 2009

By Ray "A Reader"  - See all my reviews

TOP 500 REVIEWER

Amazon Verified Purchase [\(What's this?\)](#)

This review is from: Canon PowerShot S90 10MP Digital Camera with 3.8x Wide Angle Optical Image Stabilized Zoom and 3-inch LCD (Electronics)

Early Impressions

I was delighted to have finally received my Powershot S90, and without further adieu, let me say that this is one heck of a camera. It's not perfect -- you can take truly terrible pictures with it just like you can take terrible pictures with a D700 -- but when used properly, the camera turns out remarkable shots that make us find it hard to accept the images are coming from a camera that fits in your pants pocket.

What struck me first upon using it? First, it's size. This thing is small, and it's light, too. It's a bit smaller than my Panasonic TZ3 and TZ5, and it's lighter, too. (The camera uses a front and back metal construction with plastic on the top and bottom, but the Panasonics', while also using metal, use a thicker gauge steel which adds a feel of sturdiness but also adds some weight, as well). The camera also has a high-quality feel to it. The buttons click and depress well (although the rear wheel is a bit too easy to turn, in my opinion). It has a rounded shape, so it feels comfortable in the hands, and when you stick it in your pocket, it will slide right in and out without snagging. The screen on this thing is simply gorgeous: why can't every camera have a screen like this? It's large, bright, and pretty high in resolution (461,000 pixels). You can't help but admire the camera's design once you get looking at it and using it.

Next, the camera seems to perform well in terms of speed and overall operational use. The screen has the typical lag when taking shots, but you can adjust this somewhat in the menu system to speed things up, and quite frankly, every small camera I've ever owned exhibits this behavior. It is easy to use most of the camera's functions, and you may have heard about the programmable control ring around the lens on the front of the camera. It's operation is easy, solid (the ring "clicks" with detents at different positions), and, to boot, there is the standard programmable "S" button that the Powershots' "S" cameras have typically had.

But of course, I'm interested in high ISO operation, and so I immediately took it into the livingroom where it was quite dark, and just started shooting. I was quite surprised at the results. You are not going to necessarily submit these to win any contests, but for the most part, the camera took nice shots even in that bad shooting environment, and the vast majority of the photos came out quite well (I will post a few with this review). The camera is the first (along with the Powershot G11) to deploy Sony's new ICX685CQZ sensor, a 9.31mm diagonal sensor with high performance specifications. With a little post processing, many of them look quite good. As the ISO crept into the very high ranges (800 and above) some sensor noise became apparent, but this is certainly the best low-light performance I've seen in a non-DSLR so far. (See my explanations, below, to see why this is possible). Surprisingly, some shots as high as 1600 ISO seemed to be acceptable as long as you are not a "pixel peeper." I was quite surprised when I discovered that a few of the shots had been taken at this high ISO 1600 level -- I've never had this experience before with a point and shoot camera.

Outdoor operation is fantastic. My outdoor shots for the most part have come out very well, with rich color, great detail, and little sensor noise. Like most Canon portables, these images seem to respond well to post-processing (you can sharpen them quite easily, and Canon now uses a standard meta-data tagging format that is readable by virtually all photo editor programs.) I even turned the EV down -2/3 while outside, and the sensitivity of the camera is so good that, even with this reduction in EV, my shots came out sharp and clear. Again, I will post a few shots with this review.

Aspect

- This thing is small, and it's light, too.
- Start up and turn off time is fast.
- The low light performance is best in class, period
- The one thing I don't get is the 640X480 movie mode.

Sentiment

- This thing is **small**, and it's **light**, too.
- Start up and turn off time is **fast**.
- The low light performance is **best** in class, period
- The one thing **I don't get** is the 640X480 movie mode.

Sentiment Words

- affective words: love, satisfied, disappointed
- general evaluative words: best, excellent, bad
- aspect-specific evaluative words: small, cold, long

Sentiment Words

- affective words: love, satisfied, disappointed
- general evaluative words: best, excellent, bad
- aspect-specific evaluative words: small, cold, long

This camera is small.

The LCD is small.

Beer was cold.

Pizza was cold.

The wine list is long

The wait is long.

SLDA: Sentence LDA

ASUM: Aspect Sentiment Unification Model

automatically discover aspects and the corresponding sentiments in reviews

24,184 amazon reviews

7 product categories

27,458 yelp reviews

4 cities

320 restaurants

12 sentences per review (ave)

Observation

- This thing is small, and it's light, too.
- Start up and turn off time is fast.
- The low light performance is best in class, period
- The one thing I don't get is the 640X480 movie mode.

Observation

- This thing is small, and it's light, too.
- Start up and turn off time is fast.
- The low light performance is best in class, period
- The one thing I don't get is the 640X480 movie mode.

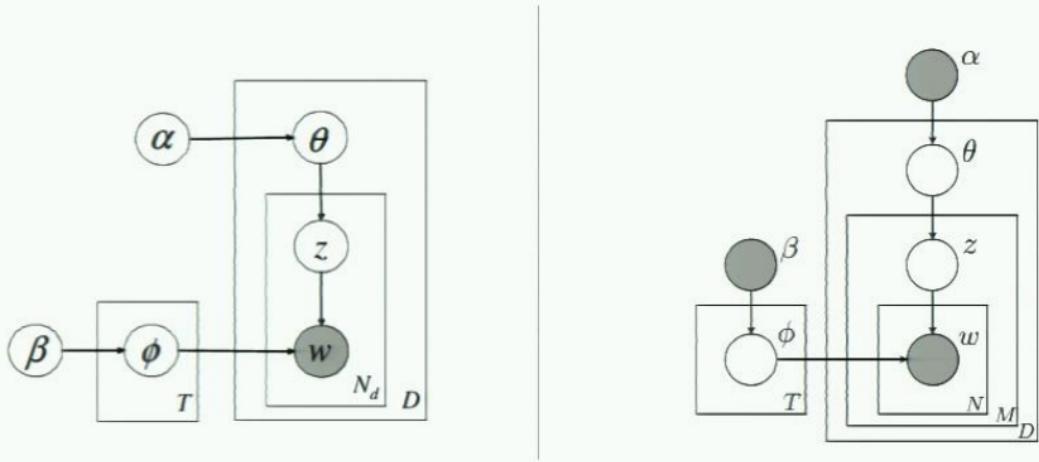
One sentence describes one aspect

Observation

- This thing is small, and it's light, too.
- Start up and turn off time is fast.
- The low light performance is best in class, period
- The one thing I don't get is the 640X480 movie mode.

One sentence describes one aspect

LDA assumption: each word represents one aspect



LDA vs SLDA

	e l e c t r o n i c s				restaurants	
camera	iso	window	keyboard	laptop	park	beer
hand	card	vista	pad	ram	street	wine
feel	raw	softwar	button	processor	valet	drink
grip	imag	mac	kei	graphic	cash	glass
weight	camera	instal	mous	netbook	lot	select
size	shoot	os	touch	drive	meter	bottl
fit	nois	xp	trackpad	core	across	martini
solid	file	run	finger	game	car	tap
small	print	program	touchpad	batteri	find	mojito
bodi	pictur	driver	scroll	hp	free	margarita

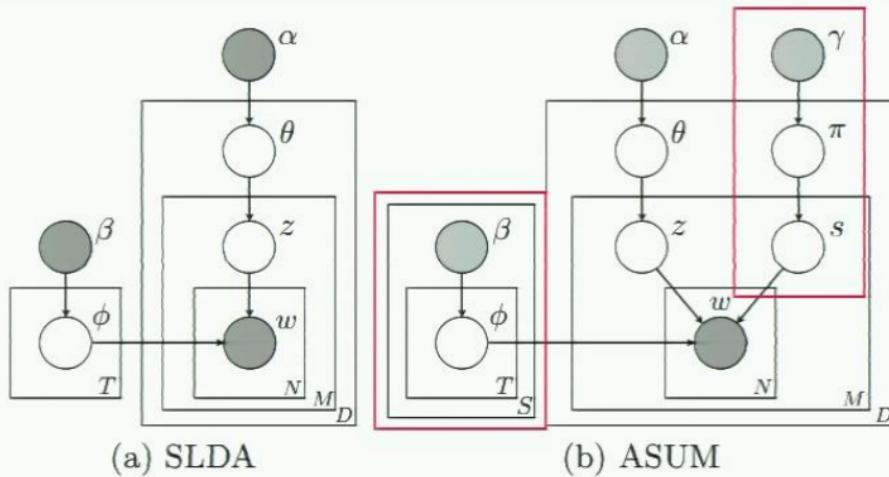
α : 0.1

β : 0.001

Aspects found by SLDA

product-specific details of reviews

Aspect-Sentiment Unification Model



topic (LDA)
aspect (SLDA)
 $\{\text{sentiment}, \text{aspect}\}$ (ASUM)

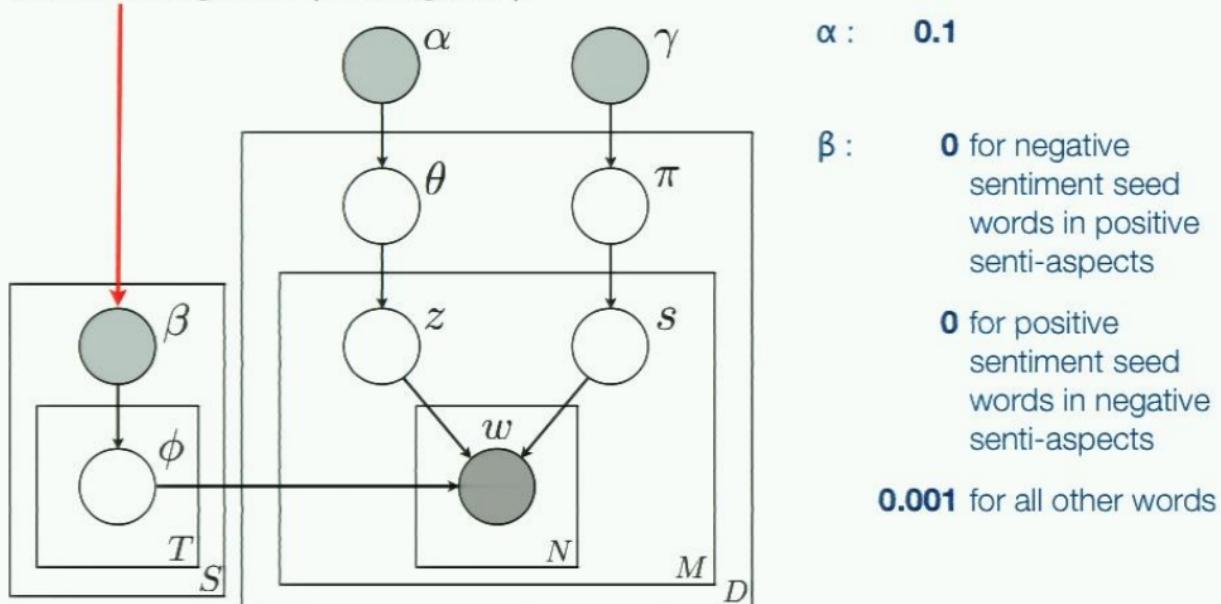
Sentiment Seed Words

PARADIGM	good, nice, excellent, positive, fortunate, correct, superior
	bad, nasty, poor, negative, unfortunate, wrong, inferior
PARADIGM+	good, nice, excellent, positive, fortunate, correct, superior, amazing, attractive, awesome, best, comfortable, enjoy, fantastic, favorite, fun, glad, great, happy, impressive, love, perfect, recommend, satisfied, thank, worth
	bad, nasty, poor, negative, unfortunate, wrong, inferior, annoying, complain, disappointed, hate, junk, mess, not_good, not_like, not_recommend, not_worth, problem, regret, sorry, terrible, trouble, unacceptable, upset, waste, worst, worthless

*built into the model by setting asymmetric priors
and Gibbs sampling initialization*

Sentiment Seed Words in the Model

β is different for positive ϕ and negative ϕ



positive senti-aspects			negative senti-aspects	
worth	screen	easi	monei	fingerprint
monei	color	light	save	glossi
penni	bright	carri	notwast	magnet
extra	clear	weight	wast	screen
well	video	lightweight	yourself	show
everi	displai	suction	notbui	finger
price	crisp	small	awai	finish
dollar	great	around	spend	print
spend	resolut	vacuum	notworth	smudg
pai	qualiti	power	stai	easili

Senti-Aspects discovered
by ASUM

contain both aspect words and
sentiment words

positive senti-aspects		negative senti-aspects		
flavor	music	dry	loud	cash
tender	night	bland	tabl	onli
crispi	group	too	convers	card
sauc	crowd	salti	hear	credit
meat	loud	tast	music	downsid
juici	bar	flavor	nois	park
soft	atmospher	meat	talk	take
perfectli	peopl	chicken	sit	accept
veri	dinner	bit	close	bring
moist	fun	littl	other	wait

Senti-Aspects discovered
by ASUM

contain both aspect words and
sentiment words

Common Words	Sentiment Words
screen color bright displai crisp qualiti sharp	clear great pictur sound movi beauti good hd imag size watch rai nice crystal
	glossi glare light reflect matt edg macbook kei black bit peopl notlik minor
music song player video download itun zune file	radio listen fm movi record easi convert podcast album audio book librari watch
	problem updat driver vista system xp firmwar disk mac hard run microsoft appl
our us server waiter tabl she he waitress ask minut seat	water glass refil wine attent friendli brought sat veri arriv plate help staff nice
	said me want card get tell if would gui bad could rude pai becaus walk then

aspect-specific sentiment
words

discovered without using sentiment
labels

I was so excited about this product.

I'd tasted the coffee and it was pretty good and easy and quick to make.
However, this machine makes the most awful, LOUD sound while heating water.

It's disturbing to hear in the morning, while others are sleeping especially!
Keurig's customer service is terrible too!

The restaurant is really pretty inside and everyone who works there looks like they like it.

The food is really great.

I would recommend any of their seafood dishes.

Come during happy hour for some great deals.

The reason they aren't getting five stars is because of their parking situation.

They technically don't "make" you use the valet but there's only a half dozen spots available to the immediate left.

senti-aspects assigned to
sentences

sentiments shown in green (p),
pink (n)

Parking (A46, Negative)

park, street, valet, lot, there, free, can, find, onli, if, valid, car, get, meter, your, block, hour, spot

- Parking is only validated for 3 hours.
- This place is a lol hard to see coming from 10th street and parking is limited.
- They don't have a lot/any designated parking/complimentary valet.
- Apparently since it's Friday the valets charge \$5 to park, which I found really annoying and just found a spot on the street.

senti-aspects assigned to
sentences

same aspect from different reviews

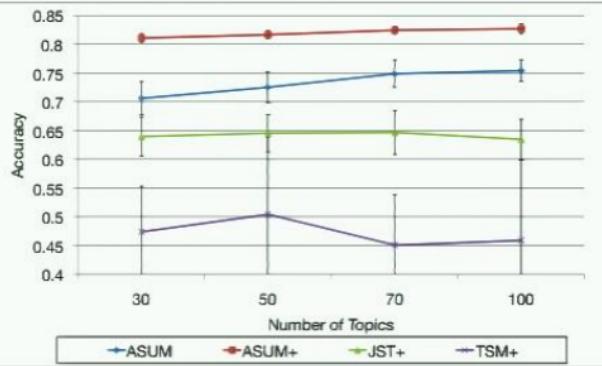
Coffeemaker Easy (A10, Positive)

coffee, hot, maker, brew, cup, great, caraf, pot, good, fast, keep, hour, love, like, machin, warm, time, thermal, easi

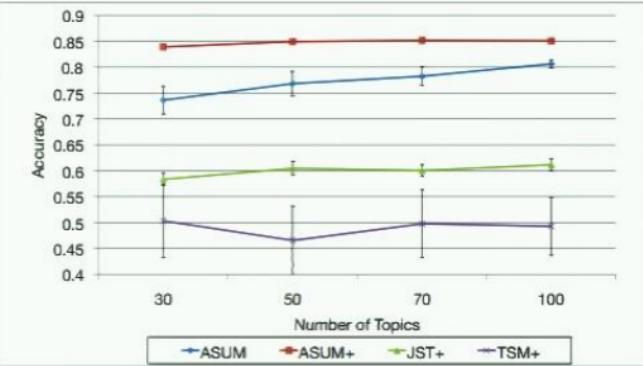
- Makes coffee fast and hot
- It took us several uses to understand how much coffee to use
- And easy to use programmer for morning coffee
- Very convenient
- Guests always comment on how nice it looks and how easy it is to use

senti-aspects assigned to
sentences

same aspect from different reviews



(a) ELECTRONICS



(b) RESTAURANTS

JST: Joint Sentiment Topic Model, Lin and He, CIKM09

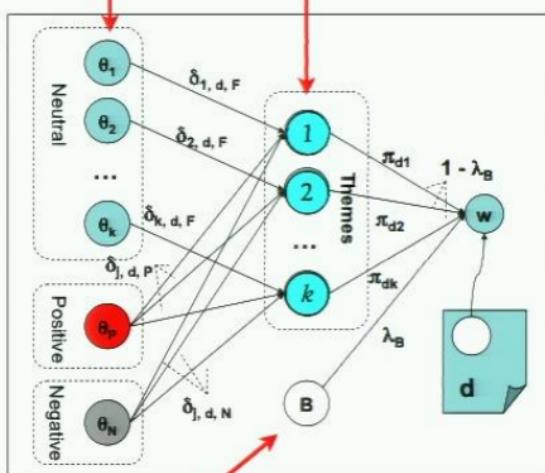
TSM: Topic Sentiment Mixture, Mei et al., WWW07

Sentiment Classification Comparison

among generative models, ASUM performs best

Language model $p(w)$ like ϕ in LDA
(Document-independent)

Like topic z in LDA
(Document-Specific)
A theme itself is not a language model



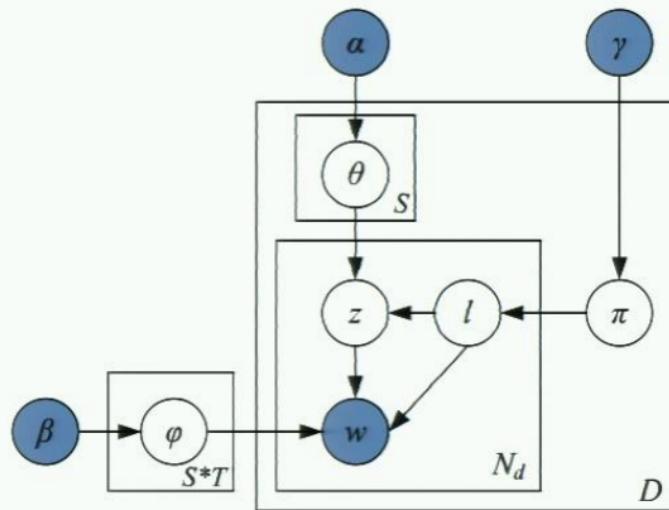
Generation Process of word w

1. Decide whether w is generated from B or themes.
If B , then choose w according to $p(w|B)$.
else
2. Choose a theme j from which w is generated.
3. Decide whether w is generated from θ_j , θ_P , or θ_N .
4. Choose w from the selected θ .

θ_P and θ_N are theme-independent (i.e., shared by all themes)

- They should cover as many sentiment words as possible to be applied to all themes
- This is problematic because it requires special effort (unlike general sentiment words)
- This model can't find theme-specific sentiment words

Language model $p(w|B)$
 B : Background words (e.g., function words)



Generation Process of word w

1. Choose a sentiment l .
2. Choose a topic label z based on l .
3. Choose w from ϕ_{zl} .

Same β for all ϕ 's

- There is no difference in β for positive ϕ and negative ϕ
- Effect of Gibbs sampling initialization fades away

There is no sentence layer required for **aspect** discovery

ASUM

Generation Process

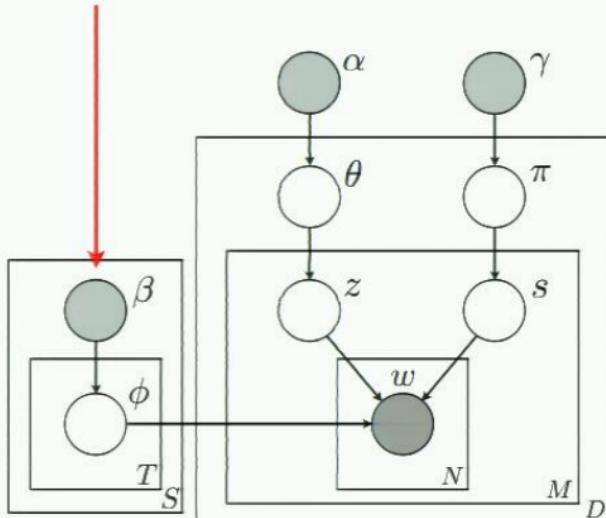
For each sentence

- Choose a sentiment s
- Choose a topic label z
- Choose words from ϕ_{zs}

Current Limitations

- θ would be different for different sentiments as in JST
- If a sentence is too short (1 or 2 words), the topic assigned is almost random (because there is no clue)
- It does not model well the sentences that have multiple aspects

β is different for positive ϕ and negative ϕ





Results on Twitter Data

- 1.3 million tweets
- 50k words in vocabulary
- What would happen when we apply this model to Twitter
 - Many more and wide variety of aspects (topics)
 - Different notions of “sentiment”
 - Review data: polarity (like vs. dislike)
 - Twitter: feelings (happy vs. unhappy)

Seed Words



:)	:()
:-)	:-('
:]	:[
:^)	:-[
:D	:('
:-D	:/
=)	:-/
=]	=('
=D	=['



Positive Senti-Aspects

vote	cream	morn	happi	idol	dinner
milei	ic	good	dai	adam	home
jona	chocol	everyon	birthdai	american	:)
cyrus	eat	:)	mother	kri	had
demi	cake	night	father	lambert	hous
award	cooki	dai	hope	allen	famili
too	butter	hello	all	vote	fun
lovato	mm	hope	thank	watch	night
teen	yum	world	mom	danc	then
taylor	peanut	twitter	bless	talent	friend
selena	yummi	afternoon	dad	paula	lunch
song	chip	all	:)	win	hang
choic	strawberri	happi	easter	susan	birthdai
who	breakfast	great	love	boyl	parti
brother	coffe	how	great	he	tonight



Positive Senti-Aspects

love	god	obama	##dollar##	app	ever
:)	bless	#tcot	dress	window	movi
thank	prai	health	sale	iphon	seen
smile	lord	palin	new	googl	funni
ya	prayer	presid	bought	instal	watch
your	jesu	vote	wear	mac	wa
keep	our	mccain	design	firefox	best
up	thank	he	shop	os	ve
ll	your	care	shoe	tweetdeck	thing
all	##time##	senat	art	chrome	funniest
we	we	republican	bag	beta	hilari
good	christ	reform	vintag	download	laugh
make	he	bill	black	us	video
lol	prais	tax	gift	version	love
alwai	love	elect	paint	desktop	saw



Negative Senti-Aspects

hurt	##percent##	tire	monei	jackson	twitter
feel	##dollar##	blah	##dollar##	michael	facebook
pain	market	:)	make	mj	try
:)	stock	sleep	guarante	rip	why
sore	price	feel	onlin	farrah	work
throat	trade	bore	earn	di	upload
headach	rate	sick	month	sad	how
sick	sale	bed	free	fawcett	can
stomach	forex	im	twitter	dead	figur
doctor	profit	sleepi	24	death	updat
flu	bank	work	home	tribut	tweetdeck
ey	report	soo	incom	billi	comput
cough	rise	realli	hour	he	link
teeth	oil	ugh	your	memori	anyon
but	billion	want	start	micheal	pictur



Negative Senti-Aspects

flight	quiz	game	obama	rain	#iranselect
airport	class	plai	health	snow	iran
plane	exam	tonight	senat	weather	avatar
drive	test	watch	state	storm	support
home	homework	laker	presid	sun	add
hour	school	footbal	new	wind	iranian
back	tomorrow	win	bill	cold	democraci
wait	done	night	court	outsid	protest
traffic	finish	readi	tax	dai	1-click
train	paper	tiger	#tcot	thunder	#iran
bu	math	let	vote	degre	green
trip	:)	wait	law	sunni	#gr88
from	work	season	fund	cloud	tehran
car	final	some	blog	but	overlai
delai	start	wing	govern	here	#twibbon

ASUM: uncovering the hidden semantic structure of aspects and sentiments

Alice Oh alice.oh@kaist.edu

Yohan Jo yohan.jo@kaist.ac.kr

<http://uilab.kaist.ac.kr>



Negative Senti-Aspects

flight	quiz	game	obama	rain	#iranelect
airport	class	plai	health	snow	iran
plane	exam	tonight	senat	weather	avatar
drive	test	watch	state	storm	support
home	homework	laker	presid	sun	add
hour	school	footbal	new	wind	iranian
back	tomorrow	win	bill	cold	democraci
wait	done	night	court	outsid	protest
traffic	finish	readi	tax	dai	1-click
train	paper	tiger	#tcot	thunder	#iran
bu	math	let	vote	degre	green
trip	:)	wait	law	sunni	#gr88
from	work	season	fund	cloud	tehran
car	final	some	blog	but	overlai
delai	start	wing	govern	here	#twibbon

ASUM: uncovering the hidden semantic structure of aspects and sentiments

Alice Oh alice.oh@kaist.edu

Yohan Jo yohan.jo@kaist.ac.kr

<http://uilab.kaist.ac.kr>



Positive Senti-Aspects

love	god	obama	##dollar##	app	ever
:)	bless	#tcot	dress	window	movi
thank	prai	health	sale	iphon	seen
smile	lord	palin	new	googl	funni
ya	prayer	presid	bought	instal	watch
your	jesu	vote	wear	mac	wa
keep	our	mccain	design	firefox	best
up	thank	he	shop	os	ve
ll	your	care	shoe	tweetdeck	thing
all	##time##	senat	art	chrome	funniest
we	we	republican	bag	beta	hilari
good	christ	reform	vintag	download	laugh
make	he	bill	black	us	video
lol	prais	tax	gift	version	love
alwai	love	elect	paint	desktop	saw