



To Swing or not to Swing: Predicting when (not) to Advertise

Donald Metzler

**Joint work with Andrei Broder, Massimiliano Ciaramita,
Marcus Fontoura, Evgeniy Gabrilovich, Vanja Josifovski,
Vanessa Murdock, and Vassilis Plachouras**

October 29, 2008
CIKM 2008



Sponsored Search

[Web](#) | [Images](#) | [Video](#) | [Local](#) | [Shopping](#) | [more](#) ▾

cheese dip recipes

Search

[Options](#) ▾



1 - 10 of about 11,000,000 for cheese dip recipes ([About this page](#)) - 0.18 sec.

Also try: [cream cheese dip recipes](#), [velveeta cheese dip recipes](#), [More...](#)

[Super Bowl Snacks](#)

[www.pacefoods.com](#) - Don't forget to serve the fans chips & Pace® Salsa for the big game.

[Spinach Cheese Dip Recipe](#)

[www.become.com](#) - Save Money. Find All That You Need. Save on What You Are Looking For.

SPONSOR RESULTS

[Cheese Dip Recipes - Food Results](#)

Popular Cuisine

[American](#) (12)

[Mexican](#) (2)

[Greek](#) (1)

[Italian](#) (1)

Top User Rated Recipes

[VELVEETA Salsa Dip](#) (Kraftfoods.com) ★★★★★ (28)

[Sweet Fruit Dip](#) (Kraftfoods.com) ★★★★☆ (15)

[Baked Triple Veggie Dip](#) (Kraftfoods.com) ★★★★☆ (24)

[Hot Parmesan Artichoke Dip](#) (Kraftfoods.com) ★★★★☆ (60)

[More...](#)

Yahoo! Shortcut - [About](#)

[Cheese Dip Recipe](#)

[Cheese Dip](#). Contributed By: kking1972. Karen King. Great ... Search [Recipes](#). Contribute [Recipe](#). Quick Meals. Meal Suggestions. Cookbook. Planning Healthy Meals ...

[www.mealsmatter.org/recipes-meals/recipe/18075](#) - 41k - [Cached](#)

SPONSOR RESULTS

[Velveeta Cheese Dip](#)

Enjoy Savings & Selection On Velveeta Cheese Dip.

[Shopzilla.com](#)

[Hidden Valley® Ranch Dip](#)

Ranch Makes Everything Better. Perfect Companion to Fresh Veggies.

[www.hiddenvalley.com](#)

[Save on Cheese Dip Recipe](#)

Give the Gift that Keeps on Giving. Shop [cheese dip recipe](#) at Exava.

[www.exava.com](#)

[Cheese Dip Recipe](#)

Make Easy & Delicious Dips! Get 1000s of [Recipes w/Recipe](#) Toolbar.

[Recipes.alot.com](#)



Content Match



With diets like this, no wonder Americans are overweight.

Skeet G - February 3, 2008 11:33:44 AM PST

Packed with fat, very high in Saturated Fat and Sodium, if you eat this, then expect to have high blood pressure, heart disease, etc, in later life.

2 of 2 found this review helpful.

Was this review helpful to you? [Yes](#) - [No](#) [Report Abuse](#)



Ditch the Velveeta

hokfota - February 3, 2008 11:12:49 AM PST

A more simple, and better tasting version of this dip is to use a brick of cream cheese and a can of chili. It's an easy one to one ratio. Microwave, and stir. It's done.

Beats Velveeta dip any day, hands down.

2 of 2 found this review helpful.

Was this review helpful to you? [Yes](#) - [No](#) [Report Abuse](#)

[View all 44 Reviews](#) | [Write a Review](#)

SPONSORED LINKS

[The Sovrana Store - Gourmet Coffee](#)

Offering the complete line of Lavazza premium coffee and espresso: whole beans or ground espresso blends, filter coffee, equipment, syrups, biscotti and more. Order online. sovranastore.com

[AHA - Culinary Education](#)

Want to be a Chef? Want to learn by "doing" instead of "watching"? www.aha.edu

[Fireside Coffee Gourmet Instant Coffee](#)

Fireside Coffee: Gourmet Flavored Instant Fat and Sugar Free Coffee. www.firesidecoffee.com

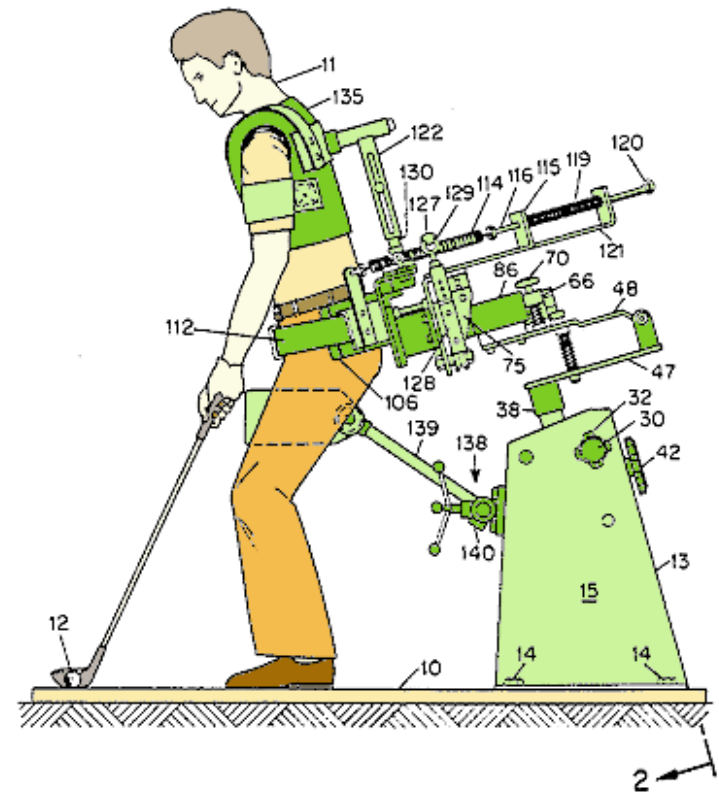
[Own Your Free Healthy Wholefood Business](#)

Wholefood Based Meals-Snacks-Wholefood Beverages-Soups-Healthy Coffee. www.capprovals.net



The Swing Problem

- Repeatedly showing non-relevant ads can have detrimental long-term effects
- Want to be able to predict when (not) to show individual ads or a set of ads (“swing”)
- Modeling actual short and long term costs of showing non-relevant ads is very difficult





Our Proposed Approaches

- Thresholding Approach
 - Decision made on individual ads
 - Only based on ad scores
- Machine Learning Approach
 - Decision made on **sets** of ads
 - Based on a variety of features



Thresholding Approach

- Set a global score threshold
- Only retrieve ads with scores above threshold
- If none of the ad scores are above the threshold, then no ads are retrieved (“no swing”)
- We ignore score normalization issues



Machine Learning Approach

- Learn a binary prediction model (“swing” or “no swing”) for an entire **set** of ads
- If we swing, then all ads are retrieved
- If we do not swing, then no ads are retrieved
- Must extract features defined over sets of ads, rather than individual ads
- Use support vector machines (SVMs)



Features

- Relevance features
 - Word overlap
 - Cosine similarity
- Vocabulary mismatch features
 - Translation models
 - Point-wise mutual information
 - Chi-squared
- Ad-based features
 - Bid price
 - Coefficient of variation of ad scores
- Result set cohesiveness features
 - Result set clarity
 - Entropy



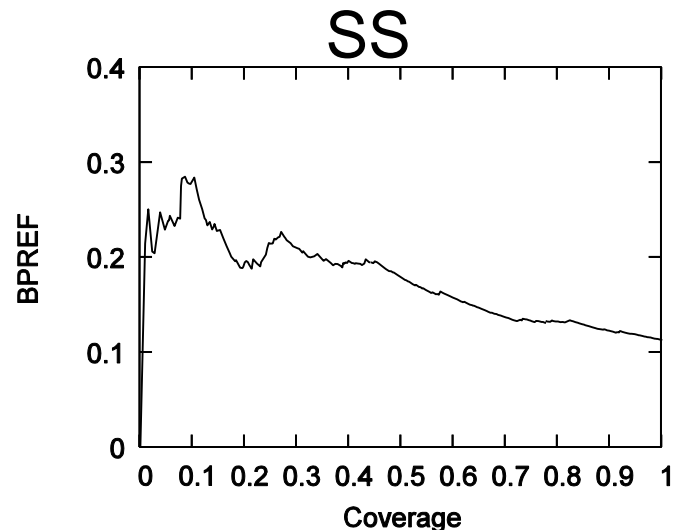
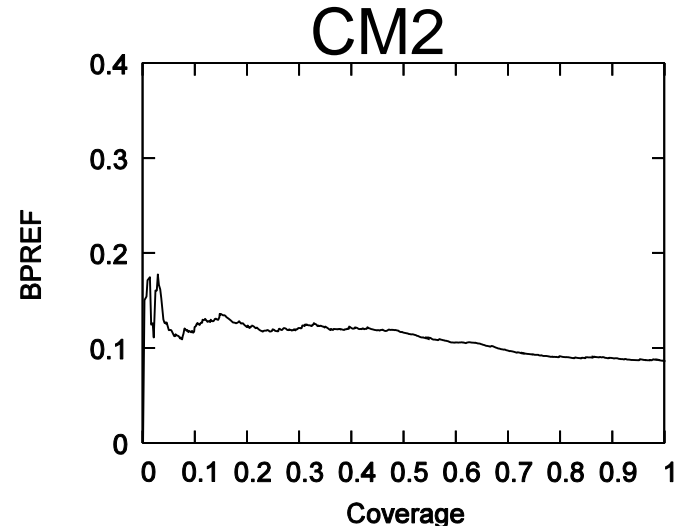
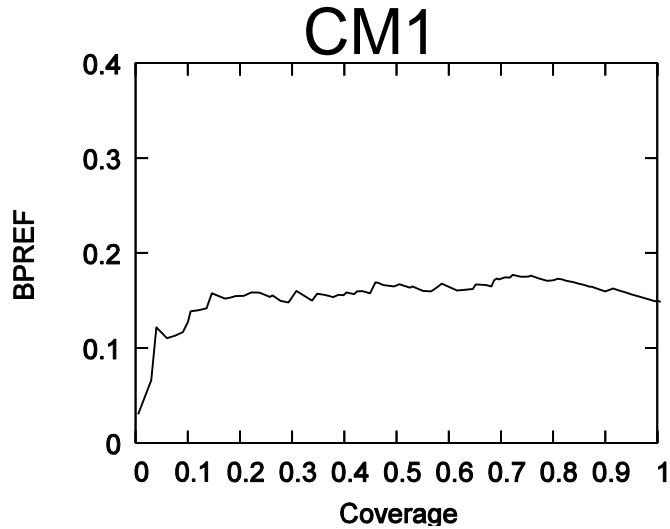
Experimental Results

- Data sets
 - Content match (CM1, CM2)
 - Sponsored search (SS)
- “Swing” / “no swing” target
 - Average all human judgments for a given query
 - If below threshold (τ), then target = “swing”

Data Set	Size	# Judgments
CM1	199 pages	5554
CM2	1103 pages	13789
SS	642 queries	8923



Thresholding Results





SVM Classification Accuracy

Data	τ	SVM Accuracy
CM1	1.4	81.90
	1.8	79.90
	2.2	78.39
	2.6	76.88
	∞	100
CM2	1.4	93.10
	1.8	80.87
	2.2	70.26
	2.6	69.27
	∞	100
SS	3.0	80.22
	3.5	69.63
	4.0	63.08
	4.5	70.09
	∞	100

- Classification accuracy for “swing” / “no swing”
- Bold indicates statistically significant increase in accuracy over majority rules classifier
- Difficult to beat baseline in edge cases due to imbalanced data



SVM Classification Accuracy with Feature Selection

Data	τ	SVM Accuracy
CM1	1.4	87.94
	1.8	87.90
	2.2	87.94
	2.6	86.43
	∞	100
CM2	1.4	94.02
	1.8	81.69
	2.2	70.53
	2.6	72.35
	∞	100
SS	3.0	82.71
	3.5	71.96
	4.0	69.63
	4.5	73.99
	∞	100

- Feature set contains many redundant and noisy features
- Used a greedy feature selection technique to overcome these issues
- Results show consistent improvements in classification accuracy



Thresholding vs. SVM Approach

Data	τ	SVM Accuracy	Query Coverage	BPREF	
				Thresh.	SVM
CM1	1.4	87.94	14.1	.1478	.2219 †
	1.8	87.90	33.7	.1497	.2019 †
	2.2	87.94	49.7	.1653	.1917†
	2.6	86.43	58.3	.1677	.1929†
	∞	100	100	.1488	.1488
CM2	1.4	94.02	1.1	.1589	.1068
	1.8	81.69	2.9	.1182	.1382
	2.2	70.53	22.5	.1330	.1804 †
	2.6	72.35	86.7	.1059	.0943
	∞	100	100	.0863	.0863
SS	3.0	82.71	4.1	.2500†	.3889†
	3.5	71.96	13.6	.2339†	.2802†
	4.0	69.63	54.9	.1650†	.1539†
	4.5	73.99	86.57	.1272	.1254
	∞	100	100	.1148	.1148

Bold BPREF = significant increase over other approach

Dagger = significant increase over “always swing” (tau = infinity)

Summary:

SVM approach tends to outperform simple thresholding, especially for small values of tau



Conclusions

- Proposed two approaches to determine when to show ads
- Thresholding approach
 - Only shows ads above some global score threshold
 - Most effective for sponsored search
- Machine learning approach
 - Predicts over entire set of ads
 - Semantic class features important for prediction
 - Effective for both sponsored search and content match



Future Work

- Apply similar model to other applications
 - Predicting quality of query rewrites
 - Predicting web search result quality
- Develop better understanding of both short-term and long-term costs involved with showing non-relevant ads