S	T	Naive	mSDA	NaiveNN	AuxNN	ADAN	MMD	FANN	DAS-EM	DAS-SE	DAS
D	B	75.20	78.50	81.10	80.80	81.70	81.05	80.30	82.00*	**82.10***	82.05*
E	B	68.85	76.15	77.95	78.00	78.55	78.65	77.25	**80.25***	77.75	80.00*
K	B	70.00	75.65	77.75	77.85	79.25	79.70	79.20	79.95	79.60	**80.05***
B	D	77.15	80.60	80.80	81.75	82.30	82.00	81.65	82.65	82.35	**82.75***
E	D	69.50	76.30	77.00	80.65	79.70	80.10	79.55	**81.40***	79.75	80.15
K	D	71.40	76.05	79.35	78.90	80.45	79.35	76.90	81.65*	**82.15***	81.40*
B	E	72.15	75.55	76.20	76.40	77.60	76.45	76.75	80.25*	75.80	**81.15***
D	E	71.65	76.00	76.60	77.55	79.70	80.20	79.25	81.40*	80.05	**81.55***
K	E	79.75	84.20	84.85	84.05	**86.85**	85.75	85.60	85.70	85.95	85.80
B	K	73.50	75.95	77.40	78.10	76.10	75.20	77.55	81.55*	79.45*	**82.25***
D	K	72.00	76.30	78.55	80.05	77.35	79.70	78.00	80.80	79.50	**81.50***
E	K	82.80	84.45	**84.95**	84.15	83.95	81.75	83.85	84.50	83.80	84.85
Average		73.66	77.98	79.38	79.85	80.29	80.00	79.65	81.84	80.68	**81.96**

Table 3: Accuracies on the Amazon benchmark. Average results over 5 runs with random initializations are reported for each neural method. * indicates that the proposed method (DAS-EM, DAS-SE, DAS) is significantly better than other baselines with $p < 0.05$ based on one-tailed unpaired t-test.

A Results on Amazon Benchmark

Most previous works (Blitzer et al., 2007; Pan et al., 2010; Glorot et al., 2011; Chen et al., 2012; Zhou et al., 2016) carried out experiments on the Amazon benchmark released by Blitzer et al. (2007). The dataset contains 4 different domains: Book (B), DVDs (D), Electronics (E), and Kitchen (K). Following their experimental settings, we consider the binary classification task to predict whether a review is positive or negative on the target domain. Each domain consists of 1000 positive and 1000 negative reviews respectively. We also allow 4000 unlabeled reviews to be used for both the source and the target domains, of which the positive and negative reviews are balanced as well, following the settings in previous works. We construct 12 cross-domain sentiment classification tasks and split the labeled data in each domain into a training set of 1600 reviews and a test set of 400 reviews. The classifier is trained on the training set of the source domain and is evaluated on the test set of the target domain. The comparison results are shown in Table 3.

B Numerical Results of Figure 1

Due to space limitation, we only show results in figures in the paper. All numerical numbers used for plotting Figure 1 are presented in Table 4. We can observe that DAS-EM, DAS-SE, and DAS all achieve substantial improvements over baseline methods under different settings.

C CNN Filter Analysis Full Results

As mentioned in Section 4.8, we conduct CNN filter analysis on NaiveNN, FANN, and DAS. For each method, we identify the top 10 most related filters for positive, negative, neutral sentiment labels respectively, and then represent each selected filter as a ranked list of trigrams with the highest activation values on it. Table 5, 6, 7 in the following pages illustrate the trigrams from the target domain (beauty) captured by the selected filters learned on E→BT for each method.

We can observe that compared to NaiveNN and FANN, DAS is able to capture a more diverse set of relevant sentiment expressions on the target domain for each sentiment label. This observation is consistent with our motivation. Since NaiveNN, FANN and other baseline methods solely train the sentiment classifier on the source domain, the learned encoder is not able to produce discriminative features on the target domain. DAS addresses this problem by refining the classifier on the target domain with semi-supervised learning, and the overall objective forces the encoder to learn feature representations that are not only domain-invariant but also discriminative on both domains.

S	T	Naive	mSDA	NaiveNN	AuxNN	ADAN	MMD	FANN	DAS-EM	DAS-SE	DAS
E	BK	49.07	55.13	58.26	60.62	63.32	60.38	59.59	66.48*	62.37	**67.12***
BT	BK	48.17	53.53	58.48	59.86	65.62	59.66	59.28	**66.78***	61.17	66.53
M	BK	45.20	49.22	57.10	60.43	62.87	60.20	57.65	69.63*	65.24*	**70.31***
BK	E	46.43	48.22	47.15	48.45	47.42	53.32	51.27	58.59*	55.15*	**58.73***
BT	E	53.63	57.32	58.77	60.98	63.13	60.53	60.62	65.71*	61.78	**66.14***
M	E	37.93	38.13	47.28	49.60	46.57	51.55	47.23	**55.88***	53.22*	55.78*
BK	BT	45.57	50.77	48.35	48.67	46.14	49.48	50.24	49.49	**54.23***	51.30*
E	BT	48.43	53.13	54.07	55.58	50.98	54.83	56.78	**61.53***	59.52*	60.76*
M	BT	39.42	39.37	47.23	48.65	44.26	48.35	48.89	47.65	**50.67***	50.66*
BK	M	43.32	47.88	47.67	48.87	51.10	53.04	52.35	55.47*	55.13*	**55.98***
E	M	41.83	47.88	50.21	51.19	50.23	51.81	52.14	58.28*	55.60*	**59.06***
BT	M	43.55	49.62	50.27	53.11	55.35	54.43	53.84	**60.95***	56.90*	60.5*
Average		45.21	49.18	52.07	53.84	53.92	54.80	54.15	59.74	57.58	**60.24**

(a) Accuracy on the small-scale dataset under setting 1

S	T	Naive	mSDA	NaiveNN	AuxNN	ADAN	MMD	FANN	DAS-EM	DAS-SE	DAS
E	BK	49.07	52.88	58.26	57.72	57.07	57.43	56.43	57.78	**58.93**	55.20
BT	BK	48.17	47.65	58.48	58.46	59.78	56.17	57.98	61.17*	60.17*	**63.32***
M	BK	45.20	48.33	57.10	58.15	58.67	57.08	57.75	58.62	58.25	**60.77***
BK	E	46.43	47.07	47.15	48.22	49.48	45.42	51.95	**54.51***	52.47*	53.92*
BT	E	53.63	55.12	58.77	59.08	59.45	60.24	58.67	61.27	**61.42**	59.83
M	E	37.93	37.40	47.28	49.43	47.00	48.72	48.92	51.28*	51.18*	**52.88***
BK	BT	45.57	49.63	48.35	47.80	47.52	45.43	49.83	53.72*	51.23*	**54.67***
E	BT	48.43	51.98	54.07	54.37	51.28	54.92	55.42	53.10	**56.43***	56.05*
M	BT	39.43	37.73	47.23	46.92	45.73	46.68	48.48	47.18	**51.57***	49.73*
BK	M	43.32	45.97	47.67	48.79	50.20	48.76	49.47	52.37*	52.68*	**53.52***
E	M	41.83	45.12	50.21	52.31	52.57	51.50	48.18	53.63*	52.25	**55.38***
BT	M	43.55	45.78	50.27	53.55	54.68	54.55	53.41	**56.24***	56.23*	56.02*
Average		45.21	47.06	52.07	52.98	52.79	52.23	53.04	55.07	55.23	**55.94**

(b) Accuracy on the small-scale dataset under setting 2

S	T	NaiveNN	ADAN	MMD	FANN	DAS-EM	DAS-SE	DAS
Y	I	53.01	55.52	54.16	54.46	55.04	56.66*	**58.54***
C	I	51.84	55.07	53.35	53.07	57.27*	55.18	**57.28***
B	I	45.85	54.64	51.40	52.39	57.31*	54.30	**58.02***
I	Y	55.46	52.57	56.52	56.30	57.92*	58.72*	**58.92***
C	Y	61.22	60.70	60.81	56.02	61.17	59.14	**61.39**
B	Y	56.86	58.42	58.77	55.72	59.94*	58.43	**61.87***
I	C	50.38	47.27	50.49	51.04	**53.46***	51.97*	53.38*
Y	C	53.87	52.53	53.12	51.86	53.48	**54.67***	**55.44***
B	C	59.48	59.91	**61.23**	60.19	59.84	59.98	59.76
I	B	50.05	46.34	47.35	48.17	48.84	**50.81**	48.84
Y	B	**54.73**	50.82	54.43	53.54	52.87	52.95	52.91
C	B	60.47	59.99	**60.52**	55.56	57.74	58.12	59.75
Average		54.43	54.48	55.18	54.02	56.24	55.91	**57.18**

(c) Macro-F1 scores on the large-scale dataset

Table 4: Performance comparison. Average results over 5 runs with random initializations are reported for each neural method. * indicates that the proposed method (DAS, DAS-EM, DAS-SE) is significantly better than other baselines with $p < 0.05$ based on one-tailed unpaired t-test.

1	2	3	4	5
best-value-at	highly-recommend-!	nars-are-amazing	beauty-store-suggested	since-i-love
good-value-at	highly-advise-!	ulta-are-fantastic	durable-machine-and	years-i-love
perfect-product-for	gogeous-absolutely-perfect	length-are-so	perfect-length-and	bonus-i-love
great-product-at	love-love-love	expected-in-perfect	great-store-on	appearance-i-love
amazing-product-∗	highly-recommend-for	setting-works-perfect	beauty-store-for	relaxing-i-love
6	7	8	9	10
store-and-am	office-setting-thanks	car-washes-!	speed-is-perfect	!-i-recommend
cleanser-and-am	locks-shimmering-color	price-in-stores	buttons-are-perfect	!-i-highly
olay-and-am	dirty-blonde-color	products-are-priced	unit-is-superb	shower-i-slather
daily-and-need	victoria-secrets-gorgeous	car-and-burning	spray-is-perfect	spots-i-needed
shower-and-noticed	dirty-pinkish-color	from-our-store	coverage-is-excellent	best-i-use

(a) NaiveNN

1	2	3	4	5
prices-my-favorite	so-nicely-!	purchase-thanks-!	feel-wonderfully-clean	are-really-cleaning
brands-my-favorite	more-affordable-price	buy-again-!	on-nicely-builds	washing-and-cleaning
very-great-stores	shampoo-a-perfect	without-hesitation-!	polish-easy-and	really-good-shampoo
great-bottle-also	an-excellent-value	buy-this-!	felt-cleanser-than	deeply-cleans-my
scent-pleasantly-floral	really-enjoy-it	discount-too-!	honestly-perfect-it	totally-moisturize-our
6	7	8	9	10
shower-or-cleaning	definitely-purchase-again	more-affordable-price	absolutely-wonderful-!	felt-cleaner-than
water-onto-my	definitely-buy-again	a-perfect-length	perfect-for-running	flat-iron-through
bleach-your-towels	perfect-for-my	an-exceptional-value	concealer-for-my	rubbed-grease-on
pump-onto-my	definitely-order-again	’ve-enjoyed-it	moisturizing-for-my	deeply-cleans-my
water-great-for	super-happy-to	pretty-decent-layer	super-glue-even	being-cleaner-after

(b) FANN

1	2	3	4	5
bath-’s-wonderful	love-fruity-sweet	feeling-smooth-radiant	cleans-thoroughly-*	excellent-everyday-lotion
all-pretty-affordable	absorb-really-nicely	love-lavender-scented	loving-this-soap	affordable-cleans-nicely
it-delivers-fabulous	shower-lather-wonderfully	am-very-grateful	bed-of-love	fantastic-base-coat
and-blends-nicely	*-smells-fantastic	love-fruity-fragrances	shower-!-*	nice-gentle-scrub
heats-quickly-love	and-clean-excellent	perfect-beautiful-shimmer	radiant-daily-moisturizer	surprisingly-safe-on
6	7	8	9	10
shower-lather-wonderfully	highly-recommend-!	excellent-fragrance-and	its-unique-smoothing	forgeous-gragrance-mist
affordable-cleans-nicely	definitely-recommend-!	fantastic-for-daytime	smooth-luxurious-texture	wonderful-bedtime-scent
peels-great-price	love-love-!	wonderfully-moisturizing-and	’s-extremely-gentle	love-essie-polish
daughter-loves-this	highly-advise-!	lathers-great-cleans	’s-affordable-combination	perfect-beautiful-shimmer
cleans-great-smells	time-advise-!	delightful-shampoo-works	absorbs-quite-well	fantastic-coverage-hydrates

(c) DAS

Table 5: Top 5 trigrams from the target domain (beauty) captured by the top 10 most **positive-sentiment-related** CNN filters learned on E→BT. $*$ denotes a padding.

1	2	3	4	5
pads-ruined-my	simply-threw-out	hours-after-trying	junk-drawer-∗	contacted-manufacturer-about
highly-disappointed-and	reviewer-pointed-out	minutes-after-rinsing	refund-time-!	minutes-not-worth
dryers-blew-my	extracts-broke-into	disappointed-after-trying	total-fake-wen	’ve-owned-this
completely-worthless-didn’t	actually-threw-out	lips-after-trying	waste-your-time	hour-unless-it
am-disappointed-and	clips-barely-keep	dry-after-shampooing	total-fail-!	results-they-claim
6	7	8	9	10
were-awful-garbage	two-failed-attempts	auto-ship-sent	refund-and-dispose	broke-don’t-fix
what-awful-garbage	a-mistake-save	am-returning-to	refund-spend-your	sent-me-expired
and-utter-waste	a-definite-return	am-unable-to	wouldn’t-recommend-!	wearing-false-eyelashes
are-absolute-garbage	a-pathetic-limp	am-pale-ghost	not-buy-dunhill	a-temporary-fix
piece-of-junk	a-total-disappointment	got-returned-and	not-worth-returning	a-disappointment-cheap

(a) NaiveNN

1	2	3	4	5
nasty-sunburn-lol	the-worse-mascaras	stale-very-unhappy	actually-hurts-your	a-return-label
bother-returning-them	it-caused-patchy	were-horrible-failures	didn’t-bother-returning	stay-away-completely
fails-miserably-at	lifeless-disaster-enter	send-this-crap	it-hurts-your	like-bug-quit
minutes-auric-needs	it-fails-miserably	were-awful-garbage	didn’t-exist-in	a-defective-brown
severely-burned-me	feel-worse-leaving	were-horribly-red	skin-horribly-after	’d-refund-the
6	7	8	9	10
worse-with-exercise	not-stink-your	it-fails-miserably	got-progressively-worse	stopped-working-for
worse-and-after	mistake-save-your	is-ineffective-apart	gave-opposite-result	uncomfortable-i-returned
unable-to-return	nothing-!-by	but-horribly-unhealthy	another-epic-fail	i-am-returning
worse-my-face	nothing-happened-!	a-pathetic-limp	got-horribly-painful	stopped-working-shortly
poorly-in-step	nothing-save-your	a-worse-job	was-downright-painful	not-waterproof-makeup

(b) FANN

1	2	3	4	5
poorly-designed-product	a-refund-spend	completely-waste-of	smells-disgusting-!	burning-rubber-stench
defective-dryer-promising	a-refund-save	of-junk-*	smells-horribly-like	began-smelling-vomit
disgusting-smelling-thing	i-regret-spending	were-awful-garbage	does-not-straighten	reaction-and-wasted
hurts-your-scalp	just-wouldn’t-spend	worthless-waste-of	’s-false-advertising	control-and-smelled
hurts-your-hair	looked-washed-out	throwing-money-away	a-disgusting-cheap	using-this-disgusting
6	7	8	9	10
super-irritating-!	got-promptly-broke	sore-and-painful	it-caused-patchy	painful-it-hurt
strong-reaction-and	after-ive-washed	is-simply-irritating	layer-hydrogenated-alcohols	unnecessary-health-risks
really-burned-and	after-several-attempts	tight-and-uncomfortable	the-harmful-uva	uncomfortable-to-wear
very-pasty-and	this-stuff-stinks	drying-and-irritating	my-severe-dark	stinging-your-eyes
super-streaky-and	again-i-threw	goopy-and-unpleasant	a-allergic-reaction	unbearable-to-wear

(c) DAS

Table 6: Top 5 trigrams from the target domain (beauty) captured by the top 10 most **negative-sentiment-related** CNN filters learned on E→BT. $*$ denotes a padding.

1	2	3	4	5
purpose-cologne-splash	okay-cord-was	hands-feet-elbows	aggressive-in-general	but-its-okay
other-hanae-mori	cocamide-dea-is	been-sealed-tight	pimples-in-general	it-moisturizes-okay
the-mavala-peeled	coily-conditioner-is	stainless-steel-blackhead	biotin-in-general	but-moisturizes-keeps
avoid-hair-pulling	flaky-dandruff-is	severely-tight-chest	dimethicone-is-terrible	but-don't-expect
cause-rashes-stinging	quickly-cord-is	thick-nasty-callouses	but-in-general	it-lathers-ok
6	7	8	9	10
pretty-damaged-from	darker-olive-complexion	doesn't-mind-pushing	producto-por-los	feeling-didn't-last
daughter-suffers-from	stronger-healthier-or	kinda-doesn't-its	unstuck-frownies-∗	curls-didn't-last
teenager-suffers-from	natural-ingredient-however	kinda-kinky-coily	they-push-∗	extra-uv-protection
tissue-damage-during	vitamin-enriched-color	okay-job-of	uva-rays-uva	garnier-fructis-curl
the-damage-on	natural-ingredients-∗	intended-purpose-that	tend-to-slip	the-mavala-after

(a) NaiveNN

1	2	3	4	5
worse-and-after	maybe-a-refund	very-disappointing-waste	my-ears-are	pretty-neutral-neither
worse-before-improving	ok-mask-i	ok-but-clean	my-neck-line	ok-so-if
unable-to-return	ok-pining-it	ok-but-will	cause-unsightly-beads	ok-during-pregnancy
unless-your-entire	ok-try-i	ok-but-didn't	my-sporadic-line	kinda-annoying-if
horrible-in-execution	ok-tho-i	ok-nothing-special	your-ear-is	ok-this-seems
6	7	8	9	10
uncomfortable-i-returned	sticky-lathers-and	some-fading-when	are-very-painful	its-also-convenient
weak-they-bend	quickly-deep-cleans	real-disappointment-the	are-less-painful	that-also-my
claimed-faulty-∗	but-elegant-bottle	especially-noticeable-after	are-a-pain	that-may-make
suffers-from-wind	beat-the-price	progressively-worse-during	about-sum-damage	that-allows-your
as-defective-∗	and-reasonably-priced	style-unfortunately-the	offered-no-pain	its-helpful-to

(b) FANN

1	2	3	4	5
'm-kinda-pale	darker-but-nope	ok-but-horrible	noticeable-i-avoid	same-result-mediocre
a-terrible-headache	gray-didn't-cover	ok-but-didn't	however-i-lean	it-caused-patchy
but-kinda-annoying	makeup-doesn't-sweat	okay-but-doesn't	but-otherwise-ok	doesn't-cause-flare
'm-kinda-mad	dark-spots-around	okay-however-it	but-im-deciding	the-harmful-uva
i-kinda-stopped	moist-but-thats	unfortunately-straight	however-i-prefer	rather-unpleasant-smell
6	7	8	9	10
kinda-annoying-if	brutal-winter-however	higher-rating-because	nothing-for-odor	but-darker-*
pretty-bad-breakage	summer-color-however	slight-burnt-rubber	kinda-recommend-this	slightly-darker-shade
my-slight-discoloration	beige-shade-however	noticeable-tan-since	not-recommend-if	somewhat-pale-affect
smells-kinda-bad	is-okay-however	somewhat-pale-affect	noticeable-but-non	but-somewhat-heavy
look-kinda-crappy	bit-greasy-however	kinda-pale-so	nothing-special-moderate	bit-dull-heavy

(c) DAS

Table 7: Top 5 trigrams from the target domain (beauty) captured by the top 10 most **neutral-sentiment-related** CNN filters learned on E→BT. ∗ denotes a padding.