

From Pain to Praise: Aspect-Based Sentiment Analysis for Norwegian Patient Feedback

Lilja Charlotte Storset, Elma Jelin, Rebecka Maria Norman
Øyvind Bjærtne, Lilja Øvrelid, Erik Velldal

May 11, 2026



Overview

1. Background
2. Aspect-Based Sentiment Analysis
3. Methodology
4. Results
5. Use-case: Demo!



Background

- **Interdisciplinary** collaboration between LTG and NIPH



Background

- **Interdisciplinary** collaboration between LTG and NIPH
- NLP not widely adopted within healthcare



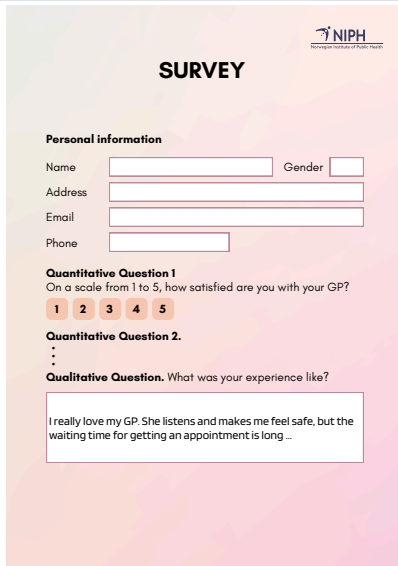
Background

- **Interdisciplinary** collaboration between LTG and NIPH
- NLP not widely adopted within healthcare
- Much unresolved potential – an important step!



Background

- NIPH conducts national patient experience surveys every year
- Several quantitative questions
- Qualitative question: free-text field



The screenshot shows a survey form from NIPH (Norwegian Institute of Public Health). The form is titled "SURVEY" and includes a section for "Personal information" with fields for Name, Gender, Address, Email, and Phone. Below this are two quantitative questions: "Quantitative Question 1" (On a scale from 1 to 5, how satisfied are you with your GP?) and "Quantitative Question 2" (On a scale from 1 to 5, how satisfied are you with your GP?). The scale is represented by five orange buttons labeled 1, 2, 3, 4, and 5. Below the quantitative questions is a qualitative question: "Qualitative Question. What was your experience like?". The answer is provided in a free-text field: "I really love my GP. She listens and makes me feel safe, but the waiting time for getting an appointment is long ...".

NIPH
Norwegian Institute of Public Health

SURVEY

Personal information

Name Gender

Address

Email

Phone

Quantitative Question 1
On a scale from 1 to 5, how satisfied are you with your GP?

1 2 3 4 5

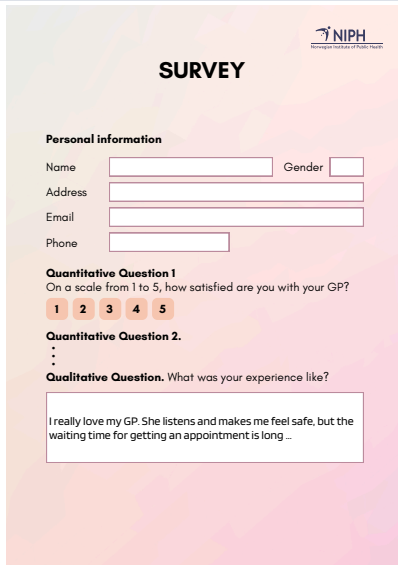
Quantitative Question 2.
:

Qualitative Question. What was your experience like?

I really love my GP. She listens and makes me feel safe, but the waiting time for getting an appointment is long ...

Background

- NIPH conducts national patient experience surveys every year
- Several quantative questions
- Qualitative question: free-text field
- Manually analyze random sample of **500 comments out of tens of thousands**



NIPH
Norwegian Institute of Public Health

SURVEY

Personal information

Name Gender

Address

Email

Phone

Quantitative Question 1
On a scale from 1 to 5, how satisfied are you with your GP?

1 2 3 4 5

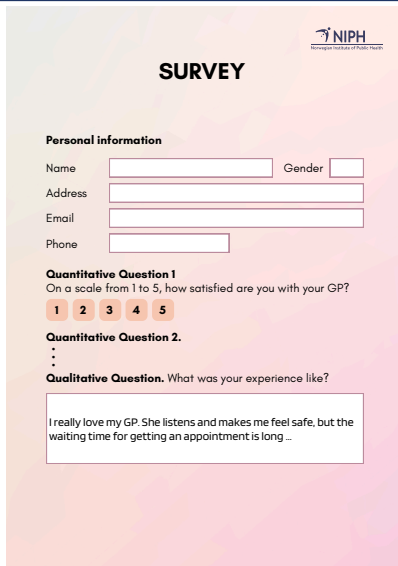
Quantitative Question 2.
⋮

Qualitative Question. What was your experience like?

I really love my GP. She listens and makes me feel safe, but the waiting time for getting an appointment is long ...

Background

- NIPH conducts national patient experience surveys every year
- Several quantitative questions
- Qualitative question: free-text field
- Manually analyze random sample of **500 comments out of tens of thousands**
- Automated analysis → all patients are heard!



NIPH
Norwegian Institute of Public Health

SURVEY

Personal information

Name Gender

Address

Email

Phone

Quantitative Question 1
On a scale from 1 to 5, how satisfied are you with your GP?

1 2 3 4 5

Quantitative Question 2.
⋮

Qualitative Question. What was your experience like?

I really love my GP. She listens and makes me feel safe, but the waiting time for getting an appointment is long ...

Aspect-Based Sentiment Analysis

ABSA focuses on determining the **sentiment** of a text with respect to a specific **aspect**.

VANSKELIG Å OPPNÅ KONTAKT MED FASTLEGEKONTORET!!!!
HARD TO GET IN TOUCH WITH THE GP OFFICE!!!!

labels: telephone and digital communication:negative



- Aspect-expanded version of the Norwegian Patient Comment corpus
- Manually annotated 8079 sentences → 2382 free-text comments
- General Practitioners (GP) & Special Mental Healthcare (SMH)
- Set of 25 fine-grained & 8 coarse-grained healthcare-related aspects

Labels

- Healthcare providers and staff
- Organization of health services
- Access and availability
- Environment and facilities
- Treatment
- Outcome and impact of treatment/stay
- Patient involvement and participation
- General



Labels

- Healthcare providers and staff
- Organization of health services
- Access and availability
 - Geographical distance to GP office
 - Telephone and digital communication
 - Waiting times in clinic
 - Waiting time for appointment
 - Workload
- Environment and facilities
- Treatment
- Outcome and impact of treatment/stay
- Patient involvement and participation
- General



Labels

- Healthcare providers and staff
- Organization of health services
- Access and availability
 - Geographical distance to GP office
 - Telephone and digital communication
 - Waiting times in clinic
 - Waiting time for appointment
 - Workload
- Environment and facilities
- Treatment
- Outcome and impact of treatment/stay
- Patient involvement and participation
- General

Sentiment labels:

positive, negative, mixed,

neutral/no aspect



Methodology

- Fine-grained & coarse-grained aspect labels
- Sentence- & comment-level

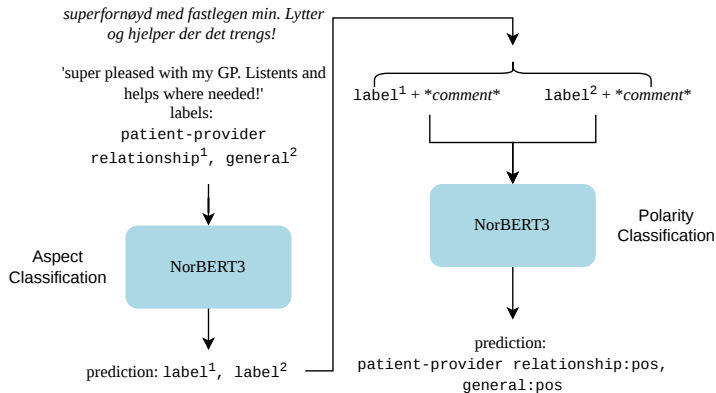


Methodology

- Fine-grained & coarse-grained aspect labels
- Sentence- & comment-level
- Encoder model: NorBERT3_large
 - Pipeline & joint approach
- Decoder model: NorMistral 11b-thinking
 - Zero-shot, few-shot, fine-tuned



NorBERT3 – Pipeline Approach



NorBERT3 – Joint Approach

Syns det er dårlig og ikke få være lenger enn ei uke!!

'Think it's too bad to not be allowed to stay longer than a week!!'

labels: duration_neg

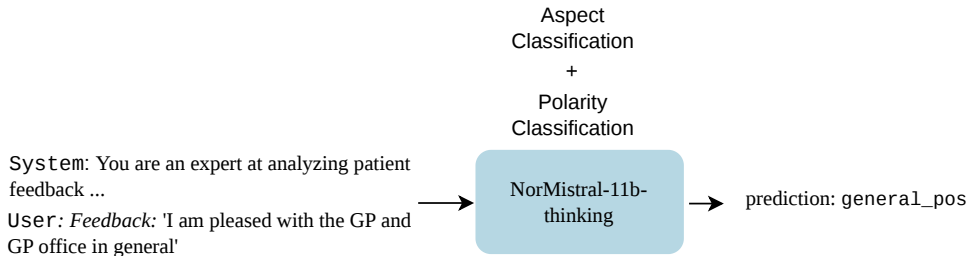
Aspect
Classification
+
Polarity
Classification

NorBERT3

prediction: duration_neg



NorMistral-11b-thinking



NorBERT3 Results

		Pipeline – Full Task		Joint – Full Task	
		GP	SMH	GP	SMH
Fine-grained	comment	60.01 $\pm$ 1.13	67.77 $\pm$ 0.98	60.75 $\pm$ 1.07	58.34 $\pm$ 0.74
	aggregated	57.17 $\pm$ 0.87	64.86 $\pm$ 0.87	60.76 $\pm$ 1.84	66.79 $\pm$ 1.66
Coarse-grained	comment	65.62 $\pm$ 0.42	67.55 $\pm$ 0.92	65.11 $\pm$ 1.94	64.44 $\pm$ 0.79
	aggregated	64.99 $\pm$ 0.6	66.56 $\pm$ 0.82	69.99 $\pm$ 1.43	70.35 $\pm$ 1.25

- Joint is better than pipeline approach – despite large label set
- Coarse-grained is better than fine-grained
 - Fine-grained not too far behind!



NorMistral Results

			GP	SMH
Zero-shot	Fine-grained	comment	27.92	32.31
		aggregated	20.87	28.8
	Coarse-grained	comment	28.66	42.91
		aggregated	28.63	37.08
Few-shot	Fine-grained	comment	35.96	34.84
		aggregated	26.65	27.22
	Coarse-grained	comment	32.62	43.86
		aggregated	30.1	37.91
Fine-tuned	Fine-grained	comment	$36.66^{\pm 2.37}$	$41.7^{\pm 1.12}$
		aggregated	$47.13^{\pm 1.93}$	$52.93^{\pm 2.5}$
	Coarse-grained	comment	$32.55^{\pm 1.4}$	$39.23^{\pm 0.99}$
		aggregated	$26.12^{\pm 1.86}$	$41.33^{\pm 3.56}$

- Encoder model is better



NorMistral Results

			GP	SMH
Zero-shot	Fine-grained	comment	27.92	32.31
		aggregated	20.87	28.8
	Coarse-grained	comment	28.66	42.91
		aggregated	28.63	37.08
Few-shot	Fine-grained	comment	35.96	34.84
		aggregated	26.65	27.22
	Coarse-grained	comment	32.62	43.86
		aggregated	30.1	37.91
Fine-tuned	Fine-grained	comment	$36.66^{\pm 2.37}$	$41.7^{\pm 1.12}$
		aggregated	$47.13^{\pm 1.93}$	$52.93^{\pm 2.5}$
	Coarse-grained	comment	$32.55^{\pm 1.4}$	$39.23^{\pm 0.99}$
		aggregated	$26.12^{\pm 1.86}$	$41.33^{\pm 3.56}$

- Encoder model is better
- Issues following the given json-format
- Over-representation of frequent labels



Demo!



Meny

Sentiment

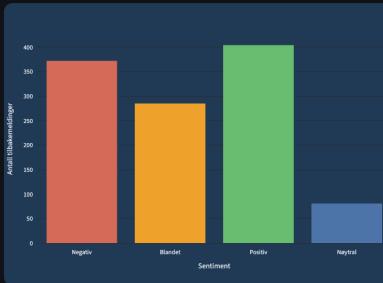
Aspekter

Kartvisning

Automatisk oppsummering

Sentimentoversikt

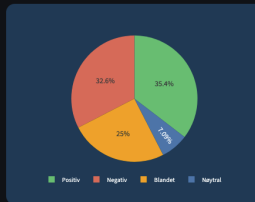
Her kan du se distribusjonen av positive, negative, blandede og nøytrale tilbakemeldinger.



Basert på

1142

Tilbakemeldinger



Meny

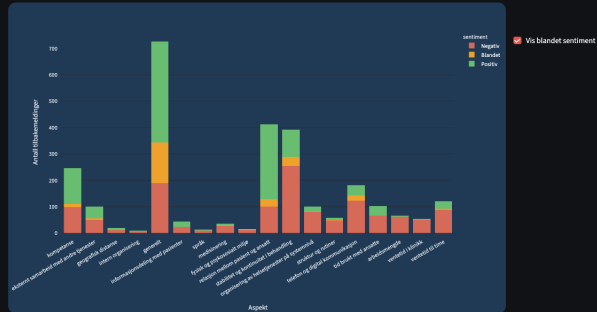
Sentiment

Aspekter

Kartvisning

Automatisk oppsummering

Oversikt alle aspekter



Oversikt per aspekt

Jeg vil se mer info om

arbeidsmengde

Arbeidsmengde

Til hvilken grad fastlegen eller andre på legekontoret virker overarbeidet og hvordan dette påvirker pasienten.

Positiv

4.5%

Negativ

95.5%

Blandet

0.0%

Basert på

66

Tilbakemeldinger

Meny

Sentiment

Aspekter

Kartvisning

Automatisk oppsummering

Kartvisning

Jeg vil se mer info om

struktur og rutiner



For fylket

Buskerud



value

0.4

0.2

0

-0.2

-0.4

Meny

Sentiment

Aspekter

Kartvisning

Automatisk oppsummering

Automatisk oppsummering

Her kan du se automatiske oppsummeringer basert på **positive** og **negative** tilbakemeldinger, med hensyn til et valgt aspekt. Sammendragene er KI-baserte. Det er kun aspekter som har 5 eller flere tilbakemeldinger som vil kunne vises.

Jeg vil se sammendrag i henhold til

ventetid til time



Pasienter fremhever generelt god tilgjengelighet til timer, spesielt hos fastlegen. Raske timer er mulig ved akutte behov, ofte via alternative leger ved samme kontor eller legevaktjenester. Fleksible timeordninger, inkludert videokonsultasjoner og "time samme dag"-prinsippet, bidrar til positiv opplevelse.

Lange ventetider er et gjennomgående problem, ofte over en måned, og noen ganger opp mot flere måneder for undersøkelser. Pasienter opplever vanskeligheter med å få time selv ved behov, og må ty til alternative løsninger som legevakt eller privatlege. Dette skyldes ofte manglende kapasitet, stor arbeidsmengde hos legene, og hyppig bruk av vikarer.

Basert på

29

Tilbakemeldinger

Basert på

75

Tilbakemeldinger

