

Report on Homework 2

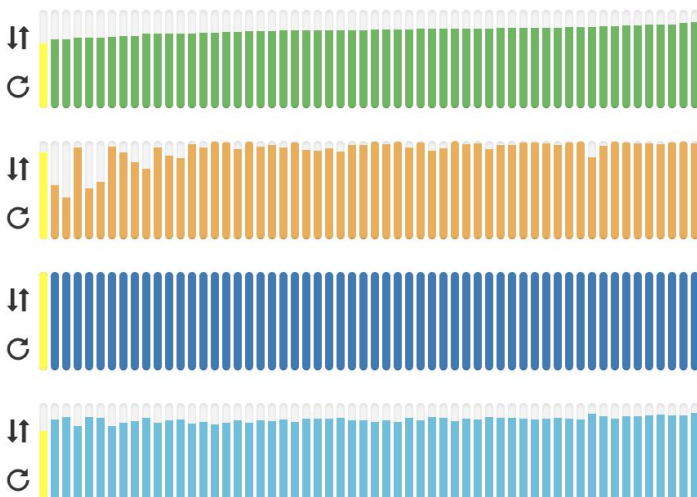
BLEU

The obtained BLEU scores:

1. accurate dev-set: 21.14
2. test-set: 25.01

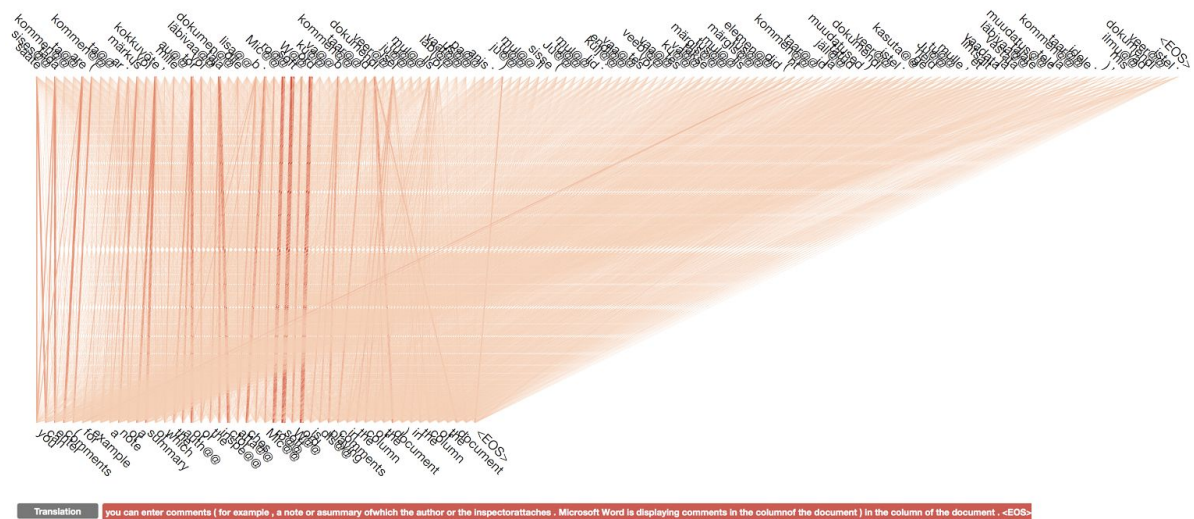
Attention analysis

There are all sorted metrics (confidence, CDP, APout, APin) for the selected sentences.

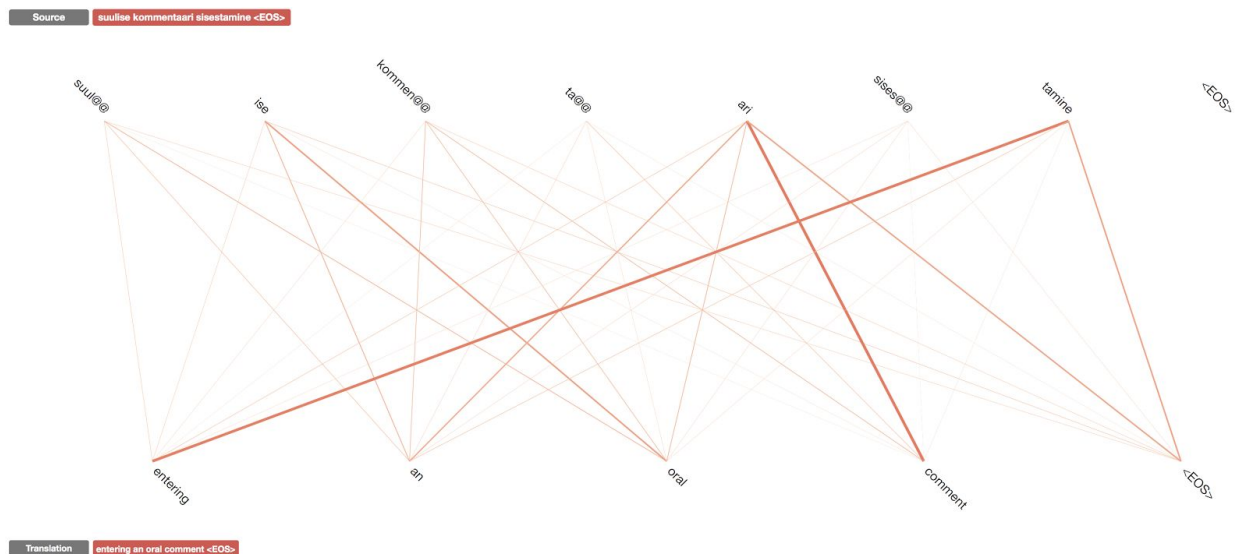


The worst confidence is provided below. We supposed that the system performed bad here because of the length of the input

sentence. We analyzed two long sentences and they both had bad BLEU scores, so this made us reach this conclusion.



The best confidence is provided below.



Sentences Analysis

Sentence 2	BLEU	Length ratio	Text
Source	-	-	Praegu puudub investeerimisosakonnal PH-funktsioon.
Human	100.00	1.00	At present the Investment department has no HRM function.
Machine	4.92	0.67	There is currently no investment department.

An assessment of the error in machine translation: From the table above we can see that machine translation faced error of the wrong meaning and cut the length of the sentence. Comparing to human translation, machine claims that there is no “investment department” at all.

Sentence 3	BLEU	Length ratio	Text
Source	-	-	Käesoleva aruande idee kujunes 22. septembri arutelus EUPCU juhi Dace Rozega.
Human	100.00	1.00	The aim of this report was formulated with Head of EUPCU Dace Roze in discussion on the 22nd of September.
Machine	17.29	1.00	The idea of this report was developed by Dace Rozega, Head of the EUPCU in the debate on 22 September.

An assessment of the error in machine translation: The meaning of the sentence was changed a bit (“idea” instead of “aim” and “developed” instead of “formulated”). Also the name was translated incorrect. In general, the system put synonyms instead of correct words from human translation..

Sentence 7	BLEU	Length ratio	Text
Source	-	-	Surve käsitleda tõhusalt ja korrektselt nii tööprotsesse kui personaliküsimusi on tungiv.
Human	100.00	1.00	The pressure to handle the working processes as well as staff questions in an effective and correct way is urgent.
Machine	2.70	0.65	Pressure to address efficiently and correctly both processes and staff matters is pressing.

An assessment of the error in machine translation: Here we can face totally wrong translation with cutted length and words' disorder in the sentence. We even can not compare to human translation word by word because of really bad quality of machine translation on this step.

Sentence 11	BLEU	Length ratio	Text
Source	-	-	Investeeringisosakonna strateegia ja lühiajaline tegevusplaan on soovitatav kord aastas läbi vaadata.
Human	100.00	1.00	It is recommendable to review the strategy and short term action plan of the Investment department yearly.
Machine	13.38	0.94	The Investment Department strategy and the short-term action plan are recommended to review the annual review.

An assessment of the error in machine translation: Here we can see that if we drop "the annual review" we would probably get better meaning representation. It looks like the system just translated one word after another.

Sentence 14	BLEU	Length ratio	Text
Source	-	-	Peagi on valmimas PRK (projekti rakendamise käsiraamat) ja PKJ (projekti korraldamise juhend).
Human	100.00	1.00	PIM (Project Implementation Manual) and POG (Project Organization Guide) are under finalising.
Machine	2.72	1.31	The PRK (project manual) and PKJ (the Code of Conduct for Project Management) will soon be ready.

An assessment of the error in machine translation: The overall flow of the machine translation variant is quite good, but it faced some problems with uppercase translation: “PRK (project manual) ” compared to “PIM (Project Implementation Manual) ” and “PKJ (the Code of Conduct for Project Management)” to “ POG (Project Organization Guide) ” respectively.

Sentence 21	BLEU	Length ratio	Text
Source	-	-	Kava taustaks on eeldus, et personalijuht/-koordinaator saab alustada täiskohaga tööd PH arendamise kallal 2003. aasta detsembri alguses.
Human	100.00	1.00	As background for the schedule is assumed that in the beginning of December 2003 the HR-Manager/Coordinator could start on full time basis to work on developing the HRM.
Machine	4.85	0.86	The plan is a prerequisite for the staff manager / coordinator to start full-time work on the development of PH in early December 2003.

An assessment of the error in machine translation: Here we can see that “The plan is a prerequisite for” is a pure machine translation compared to human one. Again faced problem with upper case translation: “HRM” compared to “PH”. As a result the cutted length error occurred.

Sentence 29	BLEU	Length ratio	Text
Source	-	-	Kord aastas tekib kaotatud tööaja tõttu kulu ka hindamisest ja PLP-vestlustest.
Human	100.00	1.00	Yearly costs arise also from the appraisal and PLP discussion in the form of lost working time.
Machine	10.57	0.94	Once a year, the cost of lost working time will also result from evaluation and PLP.

An assessment of the error in machine translation: Obviously, really poor translation with next problems occurred: the meaning was not represented well due to words' reordering; "Once a year," is grammatically correct but affected the meaning badly.

Sentence 41	BLEU	Length ratio	Text
Source	-	-	Kord aastas esitavad energiavarustusega kauplejad reguleerijale aruande teostatud ja kavandatavate tegevuste kohta ning annavad reguleerijale teavet vastavalt kommunaalteenuste reguleerijate seadusele.
Human	100.00	1.00	Once a year energy supply merchants shall submit a report to the regulator regarding accomplished and planned activity, as well as provide the regulator with information in accordance with the Law On Regulators of Public Utilities.
Machine	14.71	0.94	Once a year, energy supply traders shall submit to the regulator a report on the proposed activities and shall provide information to the regulator in accordance with the law of the Regulators' Regulators Act.

An assessment of the error in machine translation: Not bad translation but we can see missing value error as "accomplished activity" was not mentioned in machine translation variant.

Sentence 49	BLEU	Length ratio	Text
Source	-	-	1) Kauplejad, kelle omanduses või valduses on energiaülekande- või –jaotusobjektid või maagaasi hoiustamise objektid, valivad asjakohased ülekande- või jaotussüsteemi või maagaasi hoiustamissüsteemi käitajad. Maagaasivarustusega kaupleja võib asutada ühtse maagaasi ülekande-, jaotus- ja hoiustamissüsteemi käitaja.
Human	100.00	1.00	(1) Merchants in whose ownership or possession are energy transmission, distribution or natural gas storage objects shall select relevant transmission system, distribution system or natural gas storage system operators. A natural gas supply merchant may establish a unified natural gas transmission, distribution and storage system operator.
Machine	3.01	0.61	(1) Traders owned or held by energy transmission or distribution sites, select the relevant transmission or distribution system operators or operators of the storage system for natural gas.

An assessment of the error in machine translation: Here we can see that there were two sentences in Estonian as input but the MT system output is just one. In my opinion, the problem was caused by bad preprocessing. The output does not represent the meaning at all because we got “traders” in passive voice instead of the description of what they own.

Sentence 90	BLEU	Length ratio	Text
Source	-	-	Praegusest oluliselt erineva demograafilise olukorraga kohanemiseks on edaspidi vaja teha muudatusi riigi rahanduspoliitikas.
Human	100.00	1.00	Changes in public finance policy will be required in the future to adjust the budget for a substantially different demographic situation.
Machine	5.60	0.67	Changes in national monetary policy are needed to adapt to the current demographic situation.

An assessment of the error in machine translation: Here we can see that only the first part of the machine variant was correct “Changes in national monetary policy are needed to”, but the second part provides wrong meaning: “current demographic situation” compared to “substantially different demographic situation”.

Sentence 116	BLEU	Length ratio	Text
Source	-	-	1. juuli 2004. aasta seisuga seati juriidiliste isikute puhul kvalifitseeritud vabatarbija staatuse künniseks 0,0 GWh.
Human	100.00	1.00	As of July 1, 2004 the threshold for eligibility of a legal entity to the status of a qualified eligible customer was set at the 0.0 GWh.
Machine	9.16	0.59	As at 1 July 2004, the threshold for qualified customer status was set at 0,0 GWh.

An assessment of the error in machine translation: If machine translated sentence is considered separately, we can state that it can exist. But according to human translation, machine translation loses valuable part of meaning, which in some context can be crucial. Moreover, there is an grammatical mistake in date, as we can't say 'As at 1 July 2014'. Thus the use of 'at' and date format are wrong.

Sentence 128	BLEU	Length ratio	Text
Source	-	-	Iga liikmesriik otsustab, mis vormis valimised korraldatakse, aga need peavad järgima ühesuguseid demokraatlikke põhireegleid: sugude võrdõiguslikkus ja salajane hääletamine.
Human	100.00	1.00	Each Member State decides on the form its election will take, but follows identical democratic groundrules: equality of the sexes and a secret ballot.
Machine	9.19	0.96	Each Member State shall decide, in the form of elections, but must respect the same democratic basic rules: Gender equality and secret voting.

An assessment of the error in machine translation: Here we can face the change of the real meaning in the first part of the sentence “in the form of elections” compared to “on the form its election will take”. However the second part saved the overall meaning of the sentence.

Sentence 139	BLEU	Length ratio	Text
Source	-	-	Kuna Euroopa Parlamendi liikmete võim on üha suurem, mõjutavad nad oma tegevusega kõiki kodanike igapäevaelu valdkondi: keskkonda, tarbijakaitset, transporti ja ka haridust, kultuuri, tervishoidu jne.
Human	100.00	1.00	Members of the European Parliament, whose powers have become more and more extensive, influence every area of the day to day life of the European public: the environment, consumer protection and transport, as well as education, culture, health etc.
Machine	9.73	0.87	As the power of Members of the European Parliament is increasing, they are affected by all the areas of the daily lives of citizens: The environment, consumer protection, transport and education, culture, health, etc.

An assessment of the error in machine translation: Firstly we can face the wrong use of passive voice, which changed object “Members of the European Parliament” into subject by “they are affected by”. This resulted in to wrong meaning error.

Sentence 165	BLEU	Length ratio	Text
Source	-	-	Majanduses on head ajad, osalt ka seetõttu, et EL riigid lõpetavad omavahelises kaubanduses tollimaksude kogumise.
Human	100.00	1.00	It is a good period for the economy, helped by the fact that EU countries stop charging custom duties when they trade with each other.
Machine	3.39	0.60	The economy has good times, partly because the EU countries conclude customs duties on trade.

An assessment of the error in machine translation: The overall machine translation variant is very pure. “partly because” was used instead “helped by the fact that”, instead of “stop charging custom duties” was used “conclude customs duties” which is changing meaning.

Sentence 174	BLEU	Length ratio	Text
Source	-	-	1981. aastal saab EL 10. liikmeks Kreeka ning viis aastat hiljem liituvad Hispaania ja Portugal.
Human	100.00	1.00	In 1981, Greece becomes the 10th member of the EU and Spain and Portugal follow five years later.
Machine	12.91	0.89	In 1981 the EU will become Greece and five years later, Spain and Portugal will join.

An assessment of the error in machine translation: First part of the machine translation variant again change the real meaning “the EU will become Greece” compared to “Greece becomes the 10th member of the EU”, while the second part is quite good and saves the flow.

Sentence 198	BLEU	Length ratio	Text
Source	-	-	Tänu sellele teab iga üliõpilane ja muu õppur, mis on Erasmus, Leonardo ja teised piiriülese partnerluse programmid.
Human	100.00	1.00	As a result, the EU schemes for educational exchanges and trans-border partnerships like Erasmus and Leonardo are bywords among students and other learners.
Machine	4.04	0.74	Thanks to this, every student and other student are aware of Erasmus, Leonardo and other cross-border partnerships.

An assessment of the error in machine translation: One of the worst translation examples. The meaning was changed totally.

Sentence 202	BLEU	Length ratio	Text
Source	-	-	Kuna reostus ei tunne piire, on EL riigid sel alal juba aastaid ühiselt tegutsenud.
Human	100.00	1.00	As pollution knows no borders, EU member states have taken joint action in many areas.
Machine	3.23	1.13	Since pollution does not feel borders, the EU has been collectively acting in this area for years.

An assessment of the error in machine translation: Here is an example of a good translation, compared to previous variants. Surprisingly for me was that in this translation uppercase variant was translated in the right way.

Sentence 214	BLEU	Length ratio	Text
Source	-	-	Kõiki eurokupüüre ja euromünte saab kasutada kõigis riikides, kus euro on käibel.
Human	100.00	1.00	All euro notes and coins can be used in the countries where the euro is accepted.
Machine	56.81	1.00	All euro banknotes and coins can be used in all countries where the euro is turnover.

An assessment of the error in machine translation: Here we can see a perfectly represented meaning with minor word transformations. For instance, the system chosen “banknotes” instead of “notes” which in my opinion is not crucial at all. The biggest problem is that the “is accepted” was replaced by “turnover” and made the end of the sentence grammatically incorrect.

Sentence 240	BLEU	Length ratio	Text
Source	-	-	Liitumistaotluse esitamisest tegeliku liikmekssaamiseni võib riigil kuluda 10 aastat või rohkemgi.
Human	100.00	1.00	Up to 10 years or more can pass from the time a country submits an application to its actual date of entry.
Machine	7.30	0.68	When the application for membership is made, the country may spend 10 years or more.

An assessment of the error in machine translation: The meaning was changed due to using “membership is made”. This fact change the meaning flow of the sentence.

Sentenc e 248	BLEU	Length ratio	Text
Source	-	-	Kui koodikalkulaatorit välja ei lülitata, lülitub koodikalkulaator teatava ajavahemiku möödumisel automaatselt välja.
Human	100.00	1.00	In case after the use of Code calculator it has not been turned off, the Code calculator turns off automatically after a certain period of time.
Machine	4.01	0.69	If the coding calculator is not switched off, the coding calculator shall automatically switch off a specified period.

An assessment of the error in machine translation: Translation would be good if the machine used appropriate “switch off after the specified period” instead of just “switch off a specified period.”.

Sentenc e 275	BLEU	Length ratio	Text
Source	-	-	Kontaktide, kalendrite, märkmete ja veebijärjehoidjate sünkroniseerimist teie arvuti programmidega kirjeldatakse allpool.
Human	100.00	1.00	Contacts, calendars, notes, and webpage bookmarks are synced with applications on your computer, as described in the following section.
Machine	2.87	0.84	The synchronisation of contacts, calendar, notes and online synchronisation of your computer programmes are described below.

An assessment of the error in machine translation: Here machine translation system missed words like : “webpage bookmarks”. It resulted in meaning error because the list of objects were concatenated to action “and online synchronisation of”.

Sentence 347	BLEU	Length ratio	Text
Source	-	-	Pesupehmendi ja pesuvahendi kogused valige vastavalt tootja juhistele.
Human	100.00	1.00	Only use the advised quantities of fabric softener and detergent.
Machine	6.75	1.40	The quantities of detergents and detergents shall be chosen according to the manufacturer's instructions.

An assessment of the error in machine translation: Over all translation is good , but have one real drawback as machine did translate “fabric softener” as “detergents”.

Sentence 352	BLEU	Length ratio	Text
Source	-	-	Ärge mingil juhul püüdke masinat ise remontida.
Human	100.00	1.00	Under no circumstances should you attempt to repair the machine yourself.
Machine	28.42	0.82	Under no circumstances, try to repair the machine itself.

An assessment of the error in machine translation: The human translation sounds like a famous Star Wars character Yoda. The beginning of the machine translation is correct but then we can see that “should you” was dropped and “attempt” was replaced with “try”. Another point is that “yourself” appeared as “itself”. In general, the translation was simplified and the native speaker might misunderstand it totally.

Sentence 488	BLEU	Length ratio	Text
Source	-	-	Isegi traditsiooniline veebileht sureb, mille on tapnud sotsiaalsõrgustikest tulev pideva infovoo halastamatu jõud.
Human	100.00	1.00	Even the traditional website is dying, killed by the relentless force of constant streaming information from social networks.
Machine	16.02	0.83	Even traditional websites die, which has killed a continuous flow of information from social networks.

An assessment of the error in machine translation: Here we can see that machine translation lost passive voice translated “has killed a ” instead of “has been killed by a”.

Sentence 586	BLEU	Length ratio	Text
Source	-	-	Mikko Saukkonen, bändi Bonnier Amigo märgi kampaaniajuht ütles, et auhind on grupile, mis salvestab praegu teist albumit, tõmmanud juba tähelepanu.
Human	100.00	1.00	Mikko Saukkonen, promotion manager for the band's Bonnier Amigo label, said the award has already drawn attention to the group, currently recording a second album.
Machine	5.47	0.84	Mikko Saukkonen, the band Bonnier Amigo badge said that the prize is a group that recording the second album right now.

An assessment of the error in machine translation: Here we can face meaningless translation from machine because of missing words in the middle part of the sentence. However the end part of the sentence was translate good: “group that recording the second album right now.” compared to “ to the group, currently recording a second album.”. The machine variant missed “is” before “recording”.

Sentence 590	BLEU	Length ratio	Text
Source	-	-	Filmifestivalidele pole EL-i rahastatud toodang enam vöõras ja aastate jooksul on neid näinud miljonid inimesed.
Human	100.00	1.00	EU-funded productions are no strangers to film festivals and over the years have been seen by millions of people.
Machine	8.00	0.95	Film festivals are no longer funded by the EU and has seen millions of people over the years.

An assessment of the error in machine translation: Here we can see lost of the meaning because of the reordering “EU-funded productions ” compared to “Film festivals are no longer funded by the EU”.

Sentence 602	BLEU	Length ratio	Text
Source	-	-	Kaks poolt arutasid seisukohti ka detsembris Kopenhaagenis toimuva rahvusvahelise kliimamuutuse arutelu suhtes.
Human	100.00	1.00	The two sides also explored positions ahead of the international talks on climate change in Copenhagen in December.
Machine	10.71	0.72	Two sides discussed the views of the international climate change debate in Copenhagen.

An assessment of the error in machine translation: The main idea of the sentence was not changed. However as word ‘talk’ was missed, machine translation gave us such strange collocation as ‘international climate change’. Moreover, machine translation missed ‘December’ which can be influential in some cases.

Sentence 611	BLEU	Length ratio	Text
Source	-	-	See põhineb võrdsusel ja vastastikusel kasul, ühtlustades seadusandlust ning vahetades infot ja teostades praktilisi projekte.
Human	100.00	1.00	It will be based on equality and mutual benefit, by aligning the legislation and through the exchange of information and practical projects.
Machine	22.84	0.77	It is based on equality and mutual benefits, harmonising legislation and exchanging information and exercising practical projects.

An assessment of the error in machine translation: This machine translation sentence is definitely translated very good as it saves the main idea. But there are still some issues as collocation 'harmonising legislation' which sound weird, twice repeated 'and' (as first 'and' can be replaced by comma).

Sentence 637	BLEU	Length ratio	Text
Source	-	-	Läti ja Eesti nõustuvad ühiselt rentima territooriumi diplomaatilise esinduse jaoks Egiptuse Araabia Vabariigis.
Human	100.00	1.00	Latvia and Estonia agree on joint renting of premises for diplomatic representations in Arab Republic of Egypt.
Machine	30.28	1.00	Latvia and Estonia agree jointly to lease the territory of Egypt in the Arab Republic of Egypt.

An assessment of the error in machine translation: The meaning of machine translation is totally wrong, as in original sentence Latvia and Estonia agree on renting of some kind of buildings but not territories of Egypt. Also the territory of Egypt in the Arab Republic of Egypt sounds unnatural, as it seems that Egypt is included in Arab Republic of Egypt.

Sentence 643	BLEU	Length ratio	Text
Source	-	-	Lätit esindab Egiptuse Araabia Vabariigis saadik Maris Selga.
Human	100.00	1.00	Latvia is represented in the Arab Republic of Egypt by Ambassador Maris Selga.
Machine	39.65	1.23	Mr Maris Selga, who is represented in the Arab Republic of Egypt, is represented by Latvia.

An assessment of the error in machine translation: Firstly, machine translation were totally wrong to choose 'Mr' as title for Ambassador Maris Selga. Moreover the meaning of the sentence was lost as according to machine translation we can understand it as Maris Selga is represented by Latvia to Arab Republic of Egypt, which is not close to the actual meaning.

Sentence 644	BLEU	Length ratio	Text
Source	-	-	Välisminister Maris Riekstinsi telefonivestluses Venemaa välisministriga kutsus üles leidma kiiret lahendust kaubavedajate olukorrale Läti-Venemaa piiril.
Human	100.00	1.00	Foreign Minister Maris Riekstins in telephone conversation with Russian Foreign Minister invites to seek speedy solution to cargo carriers' situation on Latvia-Russia border.
Machine	15.05	1.17	The Foreign Minister Maris Riekstins calls on the Russian Foreign Minister to find a quick solution to the situation of freight carriers at the border of Russia.

An assessment of the error in machine translation: Overall, the machine translation seems to be grammatically correct. We can see that synonym was used for the word "cargo" -> "freight". The major problem is that system dropped Latvia at all which damages the meaning. According to NMT Attention Alignments we observed that "Latvia" has the strongest connection with "Russia". We suppose that this is because in training data "Latvia" appeared with "Russia" more often than alone.

Sentence 657	BLEU	Length ratio	Text
Source	-	-	Minuga saab ühendust meiliaadressil eesperenimi@gmail.com või numbril 222-222-2222.
Human	100.00	1.00	I can be reached at firstlastname@gmail.com or 222-222-2222.
Machine	4.37	1.63	I will be in contact with the privilege @ gmail.com or number 222-222-2222.

An assessment of the error in machine translation: Here we can see that the meaning was lost at all. In particular, “will be in contact” changed the meaning crucially comparing to the “can be reached”. The main reason for this might be that we don’t have sentences with similar structure in train data.

Sentence 662	BLEU	Length ratio	Text
Source	-	-	Sinu töötulemused on igakülgself ületanud meie ootusi ja me oleme õnnelikud, et saame sinust teha püsitöötaja - GlassONE meeskonna üheks osaks.
Human	100.00	1.00	Your results have exceeded our expectations in every facet of the work, and we are happy to make you a permanent employee - an integral part of the GlassONE team.
Machine	30.92	0.67	Your work results have exhausted our expectations and we are happy to make you a permanent worker - GlassONE team.

An assessment of the error in machine translation: The system changed the meaning of the initial sentence and dropped a part of sentence after the dash and the whole “in every facet of the work” clause. In particular, “exhausted” is not a synonym of “exceeded” while changing “employee” to “worker” warped the meaning only a bit.

Sentence 666	BLEU	Length ratio	Text
Source	-	-	Ma olen praegu õpilane DeLandi keskkoolis ja ma kaalun tulevase karjäärivõimalusena pangandust.
Human	100.00	1.00	I am currently a student at DeLand High School and I am considering banking as a future career path.
Machine	21.89	0.89	I'm a student at DeLand high school, and I consider the future career as a future career.

An assessment of the error in machine translation: The problem with this translation is that system considered “banking”’s meaning very similar to “future career”. Probably, it saw these two clauses in training data very often and overfitted. It also changed the tense from Present Continuous to Present Simple: “I am considering” -> “I consider”.

Sentence 681	BLEU	Length ratio	Text
Source	-	-	Probleemi lahendamiseks soovin, et kannaksite mu kontole üle rahasumma, mis nende jalanõude eest küsiti.
Human	100.00	1.00	To resolve the problem, I would like you to credit my account for the amount charged for my shoes.
Machine	9.61	1.16	In order to solve the problem, I wish you would take over the amount of money that was asked for their shoes.

An assessment of the error in machine translation: In this translation example we can see that the main part of the sentence was incorrectly translated with loss of the meaning. “their ” was used instead of “my”, “would take over ” was used, which destroys the base flow meaning line.

Sentence 688	BLEU	Length ratio	Text
Source	-	-	Mu tähelepanu on pälvinud see, et teie ettevõtte, The Cooking Store, on viimase kolme kuu jooksul hilinenud arvete maksmisega.
Human	100.00	1.00	It has come to my attention that your company, The Cooking Store has been late with paying their invoices for the past three months.
Machine	17.45	0.83	My attention is the fact that your company, The Cooking Store, is the last three months of paying late bills.

An assessment of the error in machine translation: This translation is broken in many senses. First of all, it is grammatically incorrect (wrong tense), secondly, several word reorderings caused the total loss of the meaning.

Sentence 691	BLEU	Length ratio	Text
Source	-	-	Te olete üks meie suurimaid kliente ja me hindame teiega tehtud äritehinguid.
Human	100.00	1.00	You are one of our biggest customers, and we appreciate your business.
Machine	38.35	1.33	You are one of our biggest customers and we appreciate the commercial transactions you have done.

An assessment of the error in machine translation: Over all the translation is good, but the machine decided to use “the commercial transactions” instead of using “business”, which is more appropriate in this example.

Sentence 707	BLEU	Length ratio	Text
Source	-	-	Et mõista selle arvu suurusjärku, siis 1 TeV on umbes see energiahulk, mida kasutab lendav sääsk liikumiseks.
Human	100.00	1.00	To understand the magnitude of this number, 1 TeV is the approximate amount of energy used for motion... by a mosquito in flight.
Machine	5.55	0.91	In order to understand the size of this figure, 1 TEN is about this energy, which is used by flying mosquito.

An assessment of the error in machine translation: There are two main errors here: Incorrect translation of upper case : “TEN” instead of “TeV” and machine variant missed meaningful info about energy “energy used for motion... by a mosquito in flight.” and result it into just pure “energy”. As a result we faced meaning loss in translation.

Sentence 715	BLEU	Length ratio	Text
Source	-	-	Need keskenduvad supersümmeetriliste osakeste ja Higgsi bosonite avastamisele.
Human	100.00	1.00	They will concentrate on the detection of supersymmetric particles and Higgs bosons.
Machine	26.58	0.83	They focus on the detection of supersymmetrical particles and Higgs.

An assessment of the error in machine translation: First of all, the wrong tense Present Simple instead of Future Simple. Second of all, the system did not recognise “Higgs bosons” as a collocation and dropped “bosons”. Third of all, the system picked the wrong form of the adjective “supersymmetric”.

Sentence 719	BLEU	Length ratio	Text
Source	-	-	Väljakutse seisneb selles, et antiosakeste esile kutsutud supersümmeetriat (samuti määratud s- osake) pole kunagi varem vaadeldud.
Human	100.00	1.00	The challenge is that the antiparticles termed supersymmetric (also designated sparticles) have never before been observed.
Machine	16.20	1.25	The challenge is that supersymmetrics called for by antiparticles (also assigned to a designated part) have never been seen before.

An assessment of the error in machine translation: In this translation we can see that “antiparticles” was considered to have equal meaning with “supersymmetric” which is incorrect. Again, incorrect tense in “have never been seen”.

Sentence 726	BLEU	Length ratio	Text
Source	-	-	Teades enne kokkupõrget osakeste energiad, võrdlevad nad pärast kokkupõrget tekkinud osakeste energiad.
Human	100.00	1.00	Knowing the energies of the particles prior to collision, they will be compared to the energies of the particles produced after the collisions.
Machine	2.72	0.43	Knowing before collision, they comparatively comparative energy sources after collision.

An assessment of the error in machine translation: Lots of skipped words; grammatically incorrect; vanished meaning.

Sentence 745	BLEU	Length ratio	Text
Source	-	-	Kvarkide kohta üteldakse, et nad on vangis.
Human	100.00	1.00	Quarks are said to be confined.
Machine	4.37	0.83	They say they're in jail.

An assessment of the error in machine translation: This translation is hard to comment because it is totally incorrect. This is mostly because the clause “to be confined” was considered to have the same meaning as “being in jail”. Moreover, the structure was changed crucially.

Sentence 763	BLEU	Length ratio	Text
Source	-	-	Kui elektron ja antielektron saavad kokku, hävitavad nad üksteist, mulle tulemuseks on energiapurse.
Human	100.00	1.00	When an electron and an antielectron meet, they annihilate each other, producing a burst of energy.
Machine	4.81	1.19	If the power and the antielectric power are together, they destroy each other, the result is the energy pressure.

An assessment of the error in machine translation: In this example, we can see that such terms as ‘electron’ and ‘antielectron’ was replaced by power and antielectric power, which have sense, but typically wrong. Moreover, ‘burst of energy’ was replaced by ‘energy pressure’ which is wrong. Furthermore, the part of machine translated sentence after the punctuation mark has weird structure, as if we want to say ‘the result is the energy pressure’, we will put dot before this part and then will start new sentence.

Sentence 777	BLEU	Length ratio	Text
Source	-	-	Anduri keskel on rajakamber. Selle kambri sees on miljonid pooljuhid, mis toodavad elektriimpulsi, kui osake vastu läheb.
Human	100.00	1.00	At the centre of the detector, there is a track chamber. Inside this chamber, there are millions of semiconductors that produce an electric pulse when touched by a particle.
Machine	19.12	0.83	In the center of the sensor, there are millions of semiconductors inside that chamber, producing electric pulse when a part of it goes against.

An assessment of the error in machine translation: Meaning of the sentence was partly lost because of concatenating two sentences into one and cutting the overall length. The use of “it goes against.” dot represent the actual “an electric pulse when touched by a particle.”.

Sentence 785	BLEU	Length ratio	Text
Source	-	-	Kuidas algosakeste olemust määratleda eraldatuna mehhanismidest, mis juhivad nende loomist?
Human	100.00	1.00	How can the nature of basic particles be defined beyond the mechanisms presiding over their creation?
Machine	12.53	0.88	How to define the nature of algosacents separated from mechanisms leading to their creation?

An assessment of the error in machine translation: In this case, we can see that machine translation system has created nonexistent word ‘algosacents’ which corresponds to ‘particles’ in English and to ‘algosakeste’ in Estonian. Such problem can appear because of wrong preprocessing, as system just uses Estonian word with another ending. Moreover, according to human translation, particles should be defined beyond the mechanisms, but machine translation system gives us ‘separated from mechanisms’, which is totally wrong.

Sentence 796	BLEU	Length ratio	Text
Source	-	-	Kuidas siis ühendada kõigesse sellesse üldist relatiivsusteooriat (välja käidud Einsteini poolt eelmise sajandi alguses), mis pöörleb täielikult gravitatsiooni ümber?
Human	100.00	1.00	How, then, to integrate into all this the theory of general relativity – proposed by Einstein at the beginning of the last century – which revolves entirely around gravitation?
Machine	11.54	0.83	So how do we connect all of this general relativity theory (Einstein by the beginning of the previous century), which returns completely to gravity?

An assessment of the error in machine translation: The loss of meaning was faced here mostly because of the loss of translations: “by” instead of “at”, “previous ” instead of “last”, instead of “revolves entirely around gravitation” was used “which returns completely to gravity”.

Sentence 812	BLEU	Length ratio	Text
Source	-	-	Nad seetõttu postuleerivad nähtamatu musta massi olemasolu, mis on vastastikmõjus muu mateeriaga ainult selle gravitatsiooni kaudu, ja sisaldab väga stabiilseid, seega ka väga kergeid osakesi.
Human	100.00	1.00	They therefore postulate the existence of an invisible dark matter that only interacts with the rest of matter by its gravitation, and consists of very stable, and therefore very light, particles.
Machine	3.97	0.71	They therefore possess invisible black mass, which is mutually mutually exclusive with other mathematics, and contains very stable, so very easy particles.

An assessment of the error in machine translation: Unfortunately, machine translation loses the main idea of source sentence, as it was good in translating words in sentence, but wrong usage of synonyms and wrong translation of collocations brakes sense.

Sentence 816	BLEU	Length ratio	Text
Source	-	-	Teine küsimus, mis on teadlaste jaoks problemaatiline, puudutab looduse nelja kuulast fundamentaalset vastastikmõju.
Human	100.00	1.00	Another question puzzling scientists concerns nature's four famous fundamental interactions.
Machine	3.01	1.70	The second question, which is problematic for researchers, concerns the fundamental interactions of nature in four months.

An assessment of the error in machine translation: In this example, machine translation was quite close to the meaning of the sentence, but it was wrong in the end of the sentence, as it uses word 'months' which is not response in any way to the source sentence.

Sentence 836	BLEU	Length ratio	Text
Source	-	-	Rääkige meile sellest, kuidas muuta oma äri paindlikumaks
Human	100.00	1.00	Talk to us about how to integrate resiliency into your business
Machine	9.08	0.82	Tell us how to make our business more flexible

An assessment of the error in machine translation: In this case, machine translation was totally wrong as instead of 'Talk to us' it uses 'Tell us' and 'your business' it uses 'our business'. Definitely, pairs mentioned above do not mean the same things. So it succeeded in translating only 'integrate resiliency' in 'flexible'.

Sentence 859	BLEU	Length ratio	Text
Source	-	-	Me tavaliselt teostame hindamise ja loome üksikasjaliku, tähtsuse järjekorras teekaardi ja tehnilise arhitektuuri.
Human	100.00	1.00	We typically perform an assessment and develop a detailed, prioritized roadmap and technical architecture.
Machine	12.04	1.21	We normally carry out an evaluation and create a detailed, priority order in order and technical architecture.

An assessment of the error in machine translation: Machine translation sentence here actually loses just part of sentence, which tells about development of detailed, prioritized roadmap, as it translates it in some messy collocation without any sense: 'detailed, priority order in order'.

Sentence 878	BLEU	Length ratio	Text
Source	-	-	info haldamise, mis kasvab astmeliselt
Human	100.00	1.00	Management of information that is multiplying exponentially
Machine	3.58	0.71	Information to be stepped up

An assessment of the error in machine translation: Here we can see an example where machine translation sentence is meaningless, as, firstly, it loses the part of noun: instead of 'Management of information' it uses 'Information'. Moreover, the part which has meaning of 'Exponential growth' was translated wrong totally.

Sentence 911	BLEU	Length ratio	Text
Source	-	-	Saate sisestada kommentaare (kommentaari: märkus või kokkuvõte, mille autor või läbivaataja dokumendile lisab. Microsoft Word kuvab kommentaarid dokumendi veerisel jutumullis või läbivaatuspaanis.) jutumulli sisse (jutumullid: küljendivaates või veebivaates kuvatakse märgistusmullis märgistuselemendid (nt kommentaarid ja jälitatud muudatused) dokumendi veeristel. Kasutage neid jutumulle, et lihtsalt vaadata ja vastata läbivaatajate muudatustele ja kommentaaridele.), mis ilmuvad dokumendi veeristel.
Human	100.00	1.00	You can insert a comment (comment: A note or annotation that an author or reviewer adds to a document. Microsoft Word displays the comment in a balloon in the margin of the document or in the Reviewing Pane.) inside balloons (balloons: In print layout view or Web layout view, markup balloons show markup elements, such as comments and tracked changes, in the margins of your document. Use these balloons to easily see and respond to reviewers' changes and comments.) that appear in the document margins.
Machine	1.56	0.42	You can enter comments (for example, a note or a summary of which the author or the inspector attaches. Microsoft Word is displaying comments in the column of the document) in the column of the document.

An assessment of the error in machine translation: Here we can see how the quality of the translation depends on the length of the sentence. System did not capture the dependencies between words causing the meaning to vanish.

Sentence 925	BLEU	Length ratio	Text
Source	-	-	3. Tippige kommentaari tekst kommentaari jutumulli või läbivaatuspaani.
Human	100.00	1.00	3. Type the comment text in the comment balloon or in the Reviewing Pane.
Machine	0.00	0.07	3.

An assessment of the error in machine translation: Here we can see typical situation where there is no translation, because of dot after the number. This problem is connected to preprocessing, as according to NMT Attention Alignments we can see that dot is strongly connected to end of sentence.

Sentence 928	BLEU	Length ratio	Text
Source	-	-	Suulise kommentaari sisestamine
Human	100.00	1.00	Insert a voice comment
Machine	15.97	1.00	Entering an oral comment

An assessment of the error in machine translation: In this translation, machine translation were wrong with choosing of write translation connected to synonymic sequence, as 'oral comment' is understandable but wrong in terms of collocations. Moreover, usage of 'Entering' instead of 'Enter' or 'Insert' is wrong, as in this case it is crucial to have infinitive form.

Sentence 930	BLEU	Length ratio	Text
Source	-	-	Suulised kommentaarid lisatakse kommentaari jutumullidesse heliobjektidena.
Human	100.00	1.00	Voice comments are added as sound objects inside comment balloons.
Machine	3.38	1.40	Oral comments shall be included in the comments on the sounding of sound plants.

An assessment of the error in machine translation: Firstly, we need to notice the same issue as in previous example (928) where it is important to use collocation 'Voice comment' instead of 'Oral comment'. Furthermore, the main idea of machine translated variant is lost, as 'sounding of sound plants' is just meaningless. The same is for 'Oral comments shall be included in the comments'.

Sentence 946	BLEU	Length ratio	Text
Source	-	-	1. Klõpsake menüü Läbivaatus jaotises Kommentaarid nuppu Uus kommentaar.
Human	100.00	1.00	1. On the Review tab, in the Comments group, click New Comment.
Machine	4.09	0.83	Click on the menu review in Commental Command UNew comment.

An assessment of the error in machine translation: Here we can see typical wrong translation when working with solid names of some web buttons or fields for user interaction: 'Review tab' was called as 'menu review', 'Comments group' as 'Commental Command'. Moreover, instead of New Comment we have 'UNew comment', which connected to Estonian name 'Uus kommentaar'. So in the last issue, the problem is connected to wrong preprocessing.

Sentence 977	BLEU	Length ratio	Text
Source	-	-	Kui mõelda info infrastruktuuri peale, tuleb tänapäevases keskkonnas mõelda globaalselt ja omada ülemaailmset perspektiivi.
Human	100.00	1.00	In today's environment, when thinking about Information Infrastructure, one needs to think global and have a worldwide perspective.
Machine	5.11	1.00	In the context of information infrastructure, a global and global perspective must be thought in the modern environment.

An assessment of the error in machine translation: Words' reordering vanished the meaning; wrong translation for "worldwide".

Sentence 997	BLEU	Length ratio	Text
Source	-	-	IBM-i talletuse ja andmete tootmisteenus - aitab teil vähendada projekti riske ja suurendada toote väärtust, tagades abi plaanimisel ja talletus- ja andmepõhise riistvara ja tarkvaratoodete rakendamisel või andmete uude keskkonda migreerimisel
Human	100.00	1.00	IBM Storage and Data Product Services - helping you mitigate project risk and increase product value by providing assistance with planning and implementation of storage- and data-related hardware and software products or migration of data to new environments
Machine	3.37	1.24	The production services of IBM and the production services - helps you reduce the risk of the project and increase the value of the product by ensuring assistance and the use of storage and data-based hardware and software, or the new environment for migration to the environment

An assessment of the error in machine translation: Here we can see how the quality of the translation depends on the length of the sentence. System did not capture the dependencies between words causing the meaning to vanish.