
NEURAL NETWORKS

HOMEWORK 3: TRANSFORMERS



Cătălin-Alexandru Rîpanu IA1-A
Faculty of Automatic Control and Computers
National University of Science and Technology POLITEHNICA Bucharest
E-mail: catalin.ripanu@upb.ro
Lecturer: Alexandru Sorici

12th January, 2025

ABSTRACT

This homework uses a GPT-style transformer model to generate Shakespearean-style text. The goal is to comprehend the entire training and inference pipeline for transformer models, with a focus on the learning process.

Project Structure

- **Dataset:** A collection of Shakespeare's plays from Kaggle. Minimal preprocessing is required to create a corpus suitable for tokenization.
- **Tokenization:** Two approaches are implemented:
 - Character-level tokenization.
 - Subword-level tokenization using pre-computed tokenizers like Hugging Face's GPT-2.
- **Model:** A decoder-only Transformer network with the following components:
 - Sinusoidal positional embeddings.
 - Multi-head self-attention with auto-regressive masking.
 - Feedforward layers with non-linear activations.
 - Layer normalization implemented from scratch.
- **Training Objectives:** The model is trained with a causal language model objective using CrossEntropyLoss, and the **PreLN** modality was adopted.

Dataset Statistics

- **Total number of dialogue lines:** 106,993
- **Number of unique plays:** 36

- **Number of unique characters:** 932
- **Average line length:** 38.7 characters

Lines per Play

Play	Number of Lines
Hamlet	4,069
Coriolanus	3,856
Cymbeline	3,799
Richard III	3,771
Antony and Cleopatra	3,657

Most Frequent Characters

Character	Number of Lines
GLOUCESTER	1,825
HAMLET	1,525
IAGO	1,121
FALSTAFF	1,073
KING HENRY V	1,052

Most Common Scenes

Scene	Number of Lines
Act 1, Scene 1	7,290
Act 2, Scene 1	7,061
Act 3, Scene 2	6,997
Act 5, Scene 1	6,693
Act 1, Scene 2	6,224

For preprocessing, several cleanings of text were employed:

- Remove stage directions (text in square brackets or parentheses) like 'Enter', 'Exit', 'Exeunt', 'Scene', 'Act', '[', ']', '(', ')', 'Aside', 'Within'
- Normalize various types of quotes and apostrophes
- Remove multiple spaces, but preserve line structure
- Remove leading/trailing whitespace
- **Data Structure:** Each line is represented as a DialogueLine dataclass containing the following fields:
 - **play_title:** Name of the play.
 - **character:** Speaking character's name.
 - **text:** The actual dialogue spoken by the character.
 - **act:** Act number of the play where the dialogue occurs.
 - **scene:** Scene number of the play where the dialogue occurs.

Training Loop Parameters

- **Number of Tokens:** 50,257 were used for training (Subword tokenizer), and 77 using Character level tokenizer.
- **Number of Epochs:** The model was trained for 20 epochs.
- **Batch Size:** Each batch contained 32 sequences.
- **Optimizer:** AdamW optimizer was used with the following parameters:

- $\beta_1 = 0.9, \beta_2 = 0.999$
- Weight decay: 0.01
- $\varepsilon = 10^{-8}$
- **Learning Rate Schedule:** CosineAnnealingLR, with a maximum number of iterations (Tmax) 35, and minimum learning rate (etamin) $3e-5$.
- **Gradient Clipping:** Gradients were clipped to a maximum norm of 1.0 to ensure stability.
- **Loss Function:** CrossEntropyLoss was used to train the model with a next-token prediction objective.

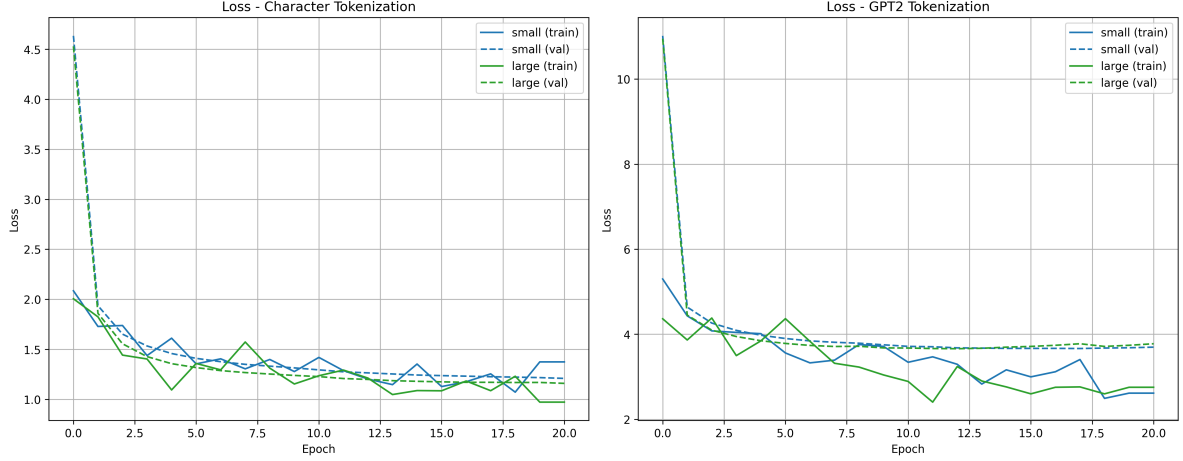


Figure 1: Comparison between loss curves

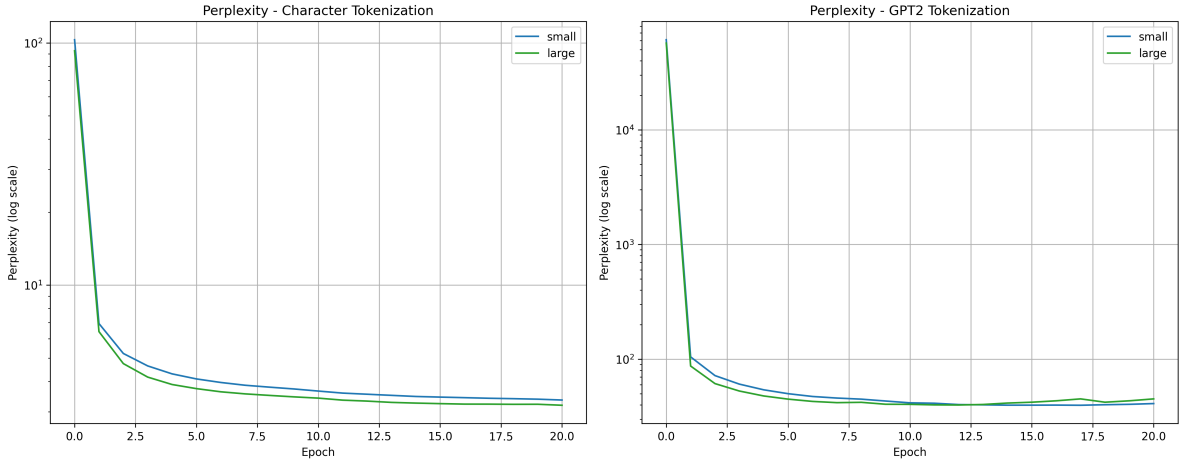


Figure 2: Comparison between perplexities

The architectural details of both models are as follows:

Parameter	Small Model	Large Model
Feature Dimension d_{model}	512	768
Number of Heads	8	12
Number of Layers	8	12
Feedforward Dimension (d_{ff})	512×4	768×4
Maximum Sequence Length	1024	1024

Table 1: Transformer Model Configurations

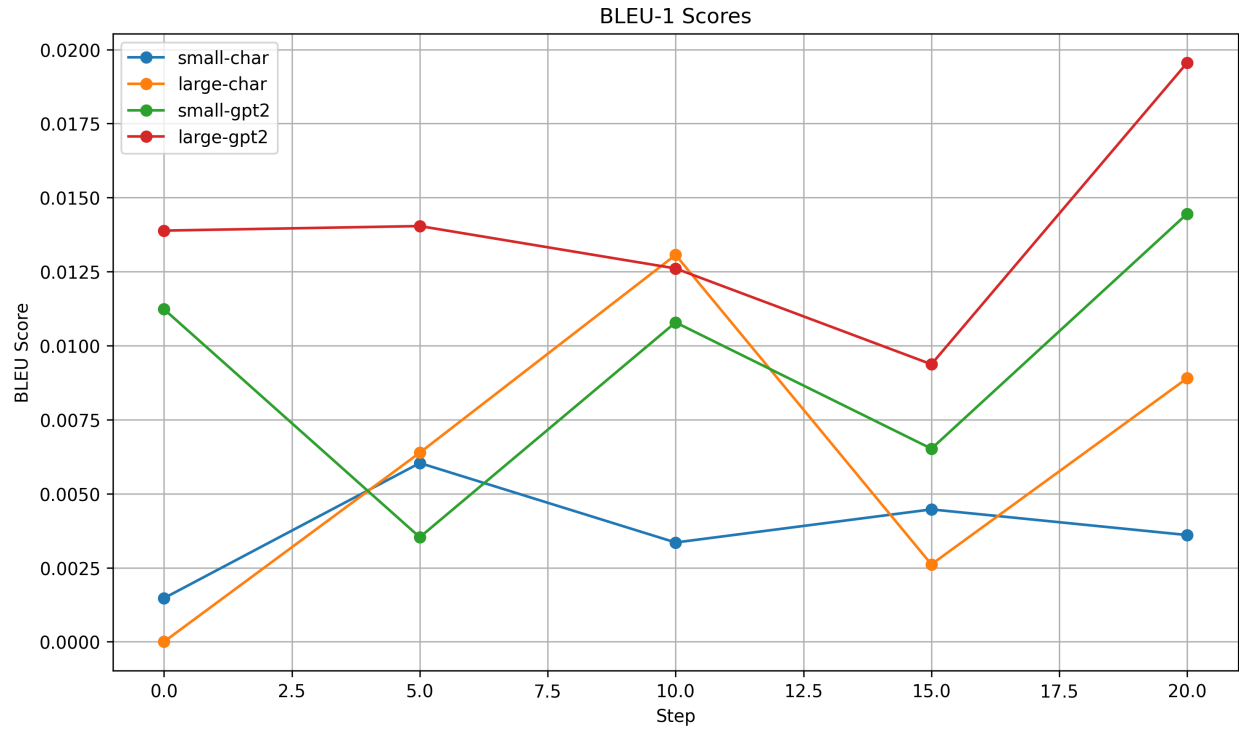


Figure 3: Comparison between BLEU scores

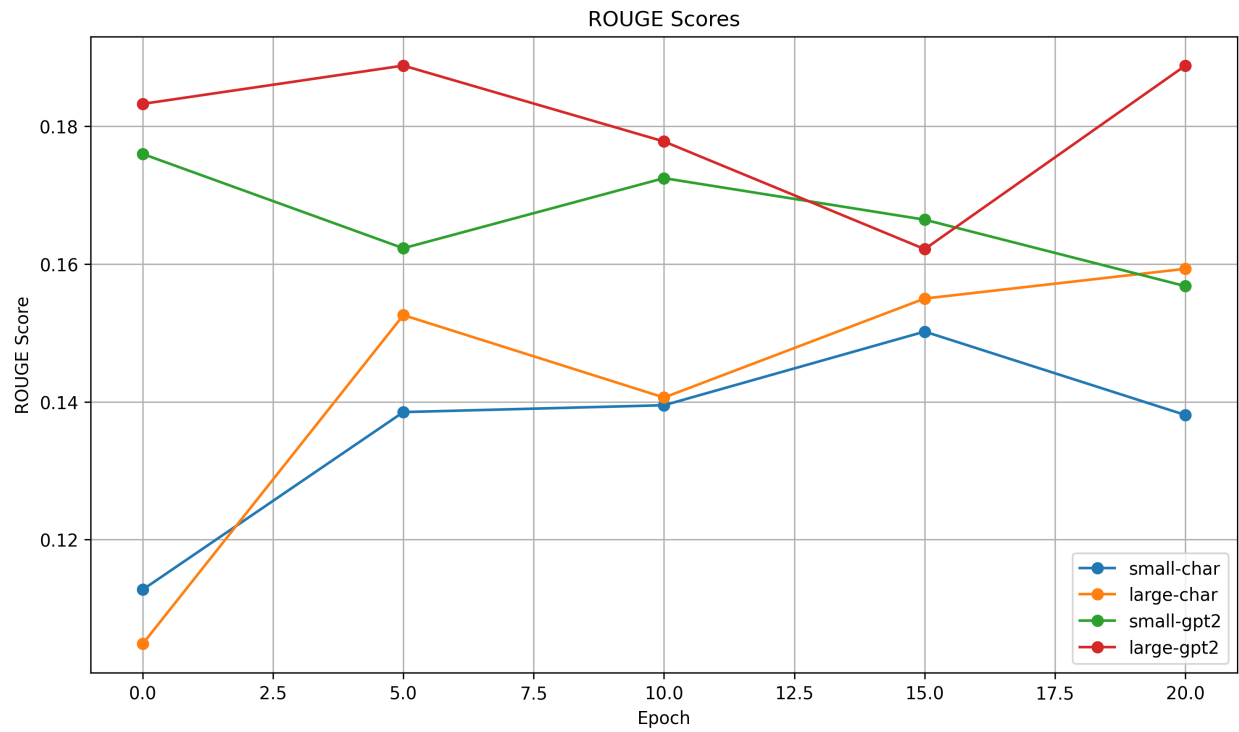


Figure 4: Comparison between ROUGE-1 scores



Figure 5: Comparison between ROUGE-2 scores



Figure 6: Comparison between ROUGE-L scores

Metric	Small - Char Tok.	Large - Char Tok.	Small - GPT-2 Tok.	Large - GPT-2 Tok.
BLEU	0.0027	0.0085	0.0148	0.0195
ROUGE-1	0.138	0.16	0.156	0.19
ROUGE-2	0.0026	0.0145	0.0125	0.0155
ROUGE-L	0.128	0.15	0.135	0.17
Perplexity	2.78	1.57	45.89	44.36
Maximum Sequence Length	1024	1024	1024	1024

Table 2: Performance Metrics for Small and Large Transformer Models with Char and GPT-2 Tokenization

As it can be seen, the best performance, in this experiment setup, was achieved by using the Large-GPT2 model. The used parameters of the top-k strategy are $k = 50$ and a temperature of 0.7 to have a balance between diversity and specific Shakespearean words.

Bonus Objectives

The following prompts given to the Large-GPT2 model are marked in bold:

HAMLET: To be, or not to be

HAMLET: Urn'd, for the time is green, but to come. It is dead. But, but one, O, when, to me, as much as it seems a king, and a fool, I should be hanged, 't, to die, it is not so, I know, you are now.

MARCELLUS: But that my lord and master is dead. What is't? speak! what's thy name? speak? I pray'd? How is't not well? How now? Is't not, my name?

GUILDENSTERN: To do you. We'll find together. We have more ready. Come. We shall. I'll to see you to the people. The other lectures them. Let them. In. I'll go with him.

FIRST PLAYER: The sun, whose eye doth give thee light, and I feel. Thou art, and dead. These, it but, and thou art not well said I, and, I, forswear thee well said I have a fool, and, I have, for the clock, for thee, I know thou art.

GERTRUDE: A song-spir-girdlemen of mine hostessant. Aumerleman, I shall not well, I am a horse, be a housewife afeithree now I have a house-a-a-shill a good fool, and a young hark, I'll be a way.

MACBETH: Is this a dagger

MACBETH: Up the same. When shall. Enter. I hope of late? Is he is my horsemen? I'll to-wain? The same. Come on. So much. Go. The same.

FORTINBRAS: And all the rest are of all reasonable ones, there's a time. Go to us. The king, my house. What is no noise o' the gate. Good, ho! Is this? GHOST: I am a gentleman, and do not think so he's dead. I have a fool, 't. The deer is coming out of wine. The devil. ' What, my gilt thou say, and a whoremaster'd, I have sent for thee, and a fool, and a fool?

LAERTES: Shall be too short, if this be a subjection but, I'll do it, do it, bear it. O, take well. That it, do it, sweet. Come, come.

HORATIO: What is the meanest thing to do? What, are you mad? I fear? no. What, my lord? What, and then tell me not? You are no more? What, my lord?

ROMEO: But, soft! what light

ROMEO: Urn is this? what's the news? what? dost thou? Is this? what dost thou? my shame to? what end? I know? what dost thou know? O friar?

FORTINBRAS: And call up my guard to the other way to the same. Come you, and make a bottle in his wedding-morrow morning to-day. What's my wife to our way to-day.

FIRST CLOWN: See you shall we shall we must. Go. The other ways. For mine armour. Come, I'll make you shall go and tell you shall go to the way to the same.

GUILDENSTERN: Not read it at supper. We will. What, my lord, to be it? We shall find you shall not. Farewell. Then, if you are at first to her. But what they are.

GHOST: Till that the sea and land have given us twain, I'll be hanged. Come, my lord. The gods! Go, let's away. , bear, go. The king, take it, I would thou shalt.

OTHELLO: O, beware, my lord, of jealousy

OTHELLO: Ills and earth, I'll speak. Good. Worthy it, O, Iago, Iago, farewell. Come, I'll bring thee a surgeon! speak it, farewell to the rest.

CLAUDIUS: What, I fear you with a cup? O, what? Let'st thou should I must be thine? Iago! and other? I perceive? what news? why should Iago, what aught?

GHOST: Go home. Farewell. My lord. Sirrah. Give me. I pray you. I am sure, I'll do you, go to-morrow thou art to-morrow with me to-enter ARman Pind you, I pray. ROSENCRANTZ: If not, I think his highness' faith is not. If he be not, he, his will be not, his letter. My lord. My lord, I will. We must be gone in his speechless hand as I'll be gone, as he is a fool, and say, I'll go to them as well met.

HAMLET: What, my boy! what's the news? What, i' faith? You? speak? Thou! What, you? what? what. Sirrah? What, Sirrah? Go hither, i' faith, my lord?

IAGO: My noble lord

IAGO: Why, what'st thou there, go with him? How now, I am I'll bring him? We'll none? What news? What news? Is this? Good-enter Fie, man?

SECOND CLOWN: For I am sure you are not a gentleman, I am king, but, and, you. O, my dear, good faith, my lord, you. What, speak, and, I know your pains, you, my lord, you love you.

GHOST: O, I am a gentleman, and a beast. O world! O most fair! a world! a woman! a woman! I! what! O slave! What! a man!

FIRST CLOWN: By heaven, you have done't. Where's Barnardine? the heavens? How now? wherefore! I pray? and these two are these? Good night? a word? O world, and a feast?

LAERTES: Go. Come, there'st me thy master sheriffetly, go. The devil. Shakest, thou shalt not seek thee, thou art thou shalt not enter I'll call him.

JULIET: O Romeo, Romeo!

JULIET: What, or I'll go to me what's-a. Come hither, boy. Come hither. Come hither, nurse, and begin to thee, I'll be gone. Come thee, I'll follow thee, I'll call.

SECOND CLOWN: A gentleman. There'st him well. When the same? I am I pray you now. The prince of his name the same? I well. Now, my lord? and I pray you at the same?

HORATIO: A compelled thing. Now, O, be it as it is! Where's the thing? what? a? O, ho! what news? or no? who? no? let'st thou, I am Iago?

MARCELLUS: And make a fool, I am a mystery by the wars. Sir John. My, I have a wife. My life with her, I were best to the same. Let me true I would have a woman of this is a cup of Prithee, I did.

FIRST PLAYER: And by some means bring thee forth to thy grave go with him go, I, go, and bear with him. Now, bear thou shalt not, speak him. Come, and take him I.

LADY MACBETH: Out, damned spot!

LADY MACBETH: Ills on both! out, friends! for shame! By my life! I see! for a cup is coming. Where'st thou suck oarsridge it is asleep. Come. Go.

FORTINBRAS: Is there no remedy? and what's the news? What is't these? What! My name? shall we? I have? and here? Dost thou meanest thou now? what colour?

GERTRUDE: I'll be a great trade about you are a great deal of the same. We'll be a song. What is a way to-a. Go to the town now, for you have a dinner you are a young Master this is a good Master Shallow her a stranger to-a-morrow we'll use you a brief, sir.

CLAUDIUS: What, I pray you? what? what trade? what, that, i' faith thy news? what. What news? My dost? what's the matter? what, what?

GUILDENSTERN: The same welcome, and go to be ready now the other men have a good, but one. I'll hear you are so much. I'll hear them all ready. I'll meet him, for you shall, by the guests.

BRUTUS: Speak, hands, for me!

BRUTUS: And get. But, and make a cup of thy guard. Come, I'll be gone. Come, and make a cup of thy ways, thou art thou see, I'll be gone.

ROSENCRANTZ: Ay, my lord. To my good lord, as much as I. They say, in the king, but I am afraid. My suit is now at large and to him I know his suit is gone from Rome, and to be gone coming forth he is true, I am so noble and bid the court to think on.

GHOST: This? Good. Go. Come. Sirrah. Good. Sirrah. The door. The same. Where is a house-morrow. Come, go. Come you next. Farewell. HORATIO: I'll give thee, my lord. It shall be well, my lord be gone. The queen, 'tis true. You are welcome. What, my lord a gentleman of good, I pray'd, sir, and I have you shall be gone.

POLONIUS: Of her that she hath made thee drink. You are gone, sir, she! I amaze me you not. , and I am a rabble-fellow of me well.

CAESAR: Et tu, Brute?

CAESAR: The same the other way to the same no more advice for this, I shall be gone. HAMLET: My tables, and my heart is coming. What is this? what, young man? what? what, are you? Prithee, Prithee what a word? is he? SECOND CLOWN: But as the performance of our England, our king, being gone, and, I'll be gone. Cade him, and fight. The queen's palace. What, and CASler, and REain. SECOND PLAYER: Pray, think thee. Now, let's not. Come hither, and let's do you, to the fellow! I'll have to him, ho, Sirrah. What is to the rest and do him, take. CLAUDIUS: And, having such a name of pity, which, being taken, , , the name, , , I protest, did see thee, did buffet not, and die, fly, and, come, and scorn'd.

FALSTAFF: What is honour?

FALSTAFF: Ills have you brought me back? I am sure you to the morning. I was out of my house. I shall. If I do that same. What's will dole or I have some other hour, I have a suit to the Lord, I have a bottle under the court?

FIRST PLAYER: The grave of a great queen. What's your will? how dost thou, my lord? nymph? O gentle lady? What, Oswald? I say you not? The dost thou, lady?

POLONIUS: Tell you I'll go to-a-aute well. The news to him, and I'll tell you with you go. Worthy your fool, you, my master I'll give you to-a-morrow pray you well.

FIRST PLAYER: Go to-morrow. Here is a young man, I'll go to-morrow. Go in. Come on, go to-day. Is there is a banquet. The clock. What woman's stay a word with me now, my lord.

ROSENCRANTZ: By my conscience, my lord, I am in the world. If I had been a brother by him, I am a brother of my heart-fellowithee. What's name of two are a soldier, let's house, he shall be true man.

BEATRICE: Against my will

BEATRICE: Come, and a good will go to-day. Come your will hold you to-morrow to-morrow to her and get a fool. Where I'll give by and swear by night he is a gentle Sir Nicholased lord.

POLONIUS: Same? the same? I shall? and doest thou? Let'st thou not the Grecians? the same? How now? let'st thou? what'st thou? what?

HAMLET: A most fair and true knight. Now, good even now, good night good night. What, ho! I pray, ho! a good morrow, good night. What, ho!

POLONIUS: If I shall I can? no? no? the matter? 'tis true, you? a dozen gold? Come. We shall? the matter? and I have we will you are you a word?

FIRST CLOWN: Thy hand my hand. What, and here go, thou shalt not meet thee, thou shalt be. Now, thou shalt not sleep. Come thee and thy voice. Now, thou shalt see, thou go to thee, there, thou shalt thou art thou shalt not alone shall find thy ways, thou art thou shalt be gone there.

KING LEAR: Blow, winds, and crack your cheeks!

KING LEAR: What! Come, sweet friend of Troy! a cup. The queen, I'll knock. What, boy! Good morrow, and here, the boy! a devil. What, I'll be my pretty!

MARCELLUS: They four foot-enter MARter at his heels, bearing two, and REine following, and REine following, and Soldiers, and safely bearing his Soldiers, and others attending arms and other Soldiers, and exit they will we shall find the other Soldiers, and exit after two, and then enter bearing the other Lords of the same to the same distance in the other c.

SECOND CLOWN: And I'll bring you to your lordship. Come, let us here. Come, my sister. Come. The king, go. Come, go. What, away. But tell him, my lord.

FIRST PLAYER: And find thee down. What noise, I shall be gone. Now, and let'st. Come, I'll go. Come on thee to the same a gate, I'll have a gate, I'll go.

GHOST: Why, what'st thou wilt thou comest thou speak? O, I have I do I do I do you speak, is with them? How now. I have. Come, what I have some of that I say, let'st thou wilt thou swear I have?

PROSPERO: We are such stuff as dreams are made on

PROSPERO: Urn, and the sea. Forres. Now, they are not, but the rabble and exitble-monster. What a city-side, and a rose-enter. Go to their, whereof, Prithee now, Pemperry.

HAMLET: For this, I know, my mother, were well resolved! O, it is the world. I, Hor, good! Ha! What, I would my lord! But, to be my lord!

GUILDENSTERN: I have done as much, I'll be hanged with her. The good will to pray you are well. Auf to danceman and tell you shall. Auf. Hereford. , I have a great master doctor of this is the court.

GHOST: You, sirrah. I'll be a tavern Biron. Go you a great house-morrow. Fare you to-a-morrow morning to dinner. Then let's go to the rabble-morrow.

FORTINBRAS: My liege, I have a heart to do it. Some say, good night. Come me. Go, Casca! A room in the city. The same long. If I have a third be gone to seek will go.

Qualitative Samples

Input Sentence 1	HORATIO: I'll cross it, though it blast me. Stay, illusion!
Prompt	HORATIO: I'll cross
Target	, though it blast me. Stay, illusion!
Small-Char Output	HORATIO: I'll crossip you,fac i'.kep,so tieu!f'-awa,kep you!saw,y'-He ifo!twa,e' siv?O' awo,fce!S' twic,se!AE awi' te.OCK,ea!AI twic,fav s'Fo!ewic,ta s'B hea!AE,fow tea!
Small-GPT2 Output	HORATIO: I'll cross your grace and my brother's daughter,–I will not go. But I'll give you leave me some more,–IUS. Well: I rather go with you to-morrow. O, sirrah: I'll go to you.
Large-Char Output	HORATIO: I'll crossle to-mig, I' to sei-have dog,gifc:ge!twa,go mic.te!sa,nu IAwе.to!ak,go mbe,siv!tco ake,div mo!ake,div mo!gou te,lie!ake mp,do twic.ie,so twa!tie,se m.
Large-GPT2 Output	HORATIO: I'll cross myself from the state of Rome. O, my lord!–O gods!–I thank you well, Othello! But–O!–Where be gone, my lord! Didst thou that?–O, my lord! and you are.

Table 3: Model Outputs for Sentence 1

Input Sentence 2	MARCELLUS: The bird of dawning singeth all night long:
Prompt	MARCELLUS: The bird
Target	of dawning singeth all night long:
Small-Char Output	MARCELLUS: The birds' coupi:wak,sav!l' to-nig.te,saw hig to-new!s' have.Two,fic!Fav he,sck!oive afic,s' he!AECI,taic!e' woup,tive!Fa how,sica!ea I'O,twick!MA he,sif owa!
Small-GPT2 Output	MARCELLUS: The bird of your heart, wherefore speak not? Othello! what, I have sent with you?–how me! O, I am not the matter. But now! what's there I heard of him,–IUS? Othe.
Large-Char Output	MARCELLUS: The birds' facp,new-mou agic,n'-to saw hig,le-spuf.ti awe!ake.ti sco!ave,tck sie.ou,ga!fo hev!tou,ab sa he!gou,tive se!oug,lvi se!oug,rab s'-euf,iou hac.te,if.
Large-GPT2 Output	MARCELLUS: The bird is not like the state for his true. If it be not, 'tis a friend: I will pass out of man. The king, 'A: he will not meet him. Come, and the matter: I'll be not so.

Table 4: Model Outputs for Sentence 2

One explanation for the GPT2-Tokenization's high perplexity is that because of the old English words and formulations in the text corpus, the sub-word technique can't enhance the generation performance by decoding the tokens due to unsuitable vocabulary.

Report

The final report includes:

- Data organization and preprocessing.
- Model architecture details (small and large variants).
- Training configurations (epochs, optimizer, learning rate, etc.).

- Loss and perplexity graphs for all experiments.
- BLEU/ROUGE score comparisons.
- Qualitative text samples and generated outputs.

References

- Kaggle Dataset: <https://www.kaggle.com/datasets/kingburrito666/shakespeare-plays>
- Hugging Face Tokenizer: <https://huggingface.co/docs/transformers>
- Reference Paper: <https://arxiv.org/pdf/2002.04745.pdf>