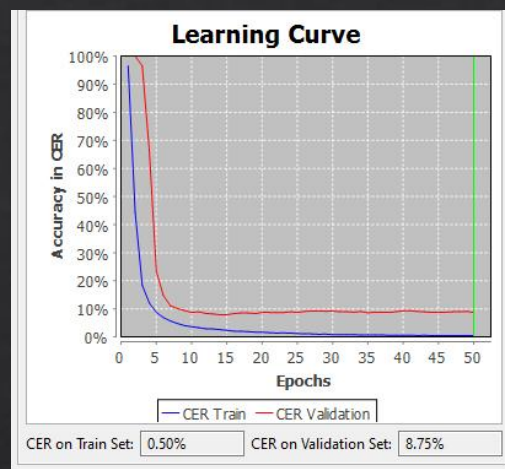


Automatická transkripčia
historických prameňov
obsahujúcich viac rukopisov na
príklade kanonických vizitácií

Prvá etapa – Vytvorenie referenčných modelov

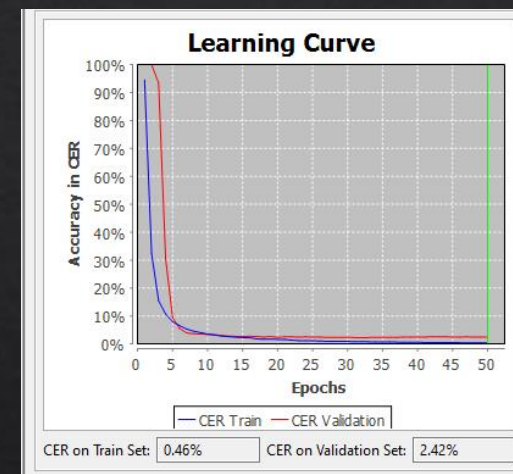
◆ Referenčný model A

- ◆ Vizitácia farnosti Špania Dolina
- ◆ Chybovosť v overovacom súbore – **8,75%**
- ◆ Veľkosť cvičnej vzorky – **2168 slov**



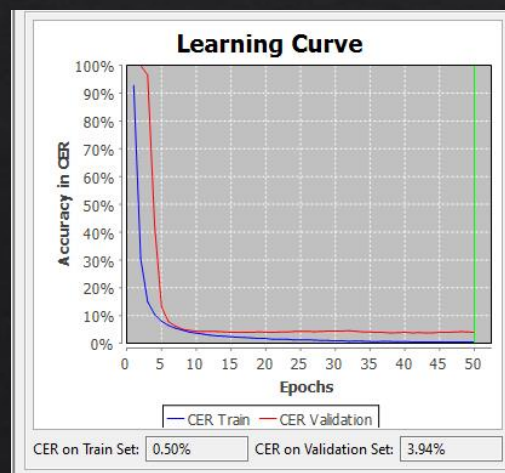
◆ Referenčný model B

- ◆ Vizitácia farnosti Staré Hory
- ◆ Chybovosť v overovacom súbore – **2,42%**
- ◆ Veľkosť cvičnej vzorky – **2991 slov**



◆ Referenčný model C

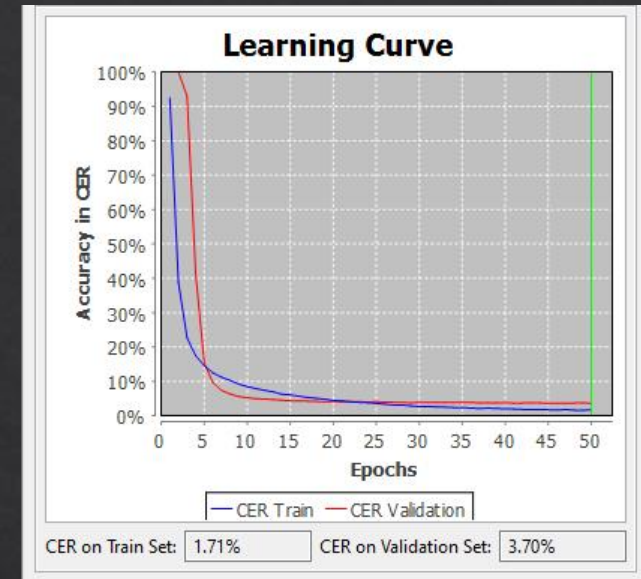
- ◆ Vizitácia farnosti Motyčky
- ◆ Chybovosť v overovacom súbore – **3,94%**
- ◆ Veľkosť cvičnej vzorky – **3232 slov**



Druhá etapa – Vytvorenie kombinovaného modelu

◇ Kombinovaný model ABC

- ◇ Založený na troch paleograficky podobných rukopisoch
- ◇ Hodnota chybovosti v overovacom súbore – 3,70%
- ◇ Veľkosť cvičnej vzorky – 8391 slov



non existit hic ulla Ecclesia, aut ~~sacellum~~. Sacellum.
4to
Reperiuntur in ea animæ ~~26~~. 26. Videlicet Confessionis
Capaces 14. Incapaces Confessionis 12. Idioma Loqvuntur
slavicum.
5to
Populus est ~~Devotus~~. Devotus, in Ecclesia freqvens Parochi reverens
bene instructus Rudimentis ~~Fidei~~. Fidei & Romano Catholicus.
6to
Diversæ Religionis Conjugati nulli dantur.
Qvarta Filialis est Tureczka Inferior. Habitationes
Laboratorum Montanorum.
Distat a Matre solida media hora Via incomoda Lapidosa
et ~~Contra~~. Contra montem.
3tio
non existit hic ulla Ecclesia aut ~~sacellum~~. Sacellum.

Tretia etapa – Využitie base modelu

Využívaný model: *NeoLatin_Ravenstein_1643-1772*

- ◇ Aplikovanie base modelu na **Referenčný model A**
 - ◇ **Zlepšenie** hodnoty chybovosti z **8,75%** na **5,50%**
- ◇ Aplikovanie base modelu na **Kombinovaný model ABC**
 - ◇ **Zlepšenie** hodnoty chybovosti z **3,70%** na **3,25%**
- ◇ Efektívnosť využitia tejto metódy závisí na správnom výbere base modelu
 - ◇ Base model by mal byť vytrénovaný na paleograficky podobnom rukopise
 - ◇ Jazyk dokumentov by mal byť rovnaký

Details

Name: NeoLatin_Ravenstein_1643-1772 Language: Latin

Description: This model is based on the transcription of the "Litterae Annuae Parochiae Ravensteijn SJ ab Anno 1643 ad Annum 1772". The annual letters were kept at the Archivum Neerlandicum Societatis Iesu (Berchmannianum, Nijmegen). These are now at the Catholic Document Center (Katholiek Documentatie Centrum (KADOC)) in Leuven, Belgium (inventarisnummer 15.606).

Parameters:

Document Type: Handwritten Show advanced parameters...

Nr. of Words: 64435 Nr. of Lines: 6864

Save Show Train Set Show Validation Set Show Characters

Learning Curve

CER on Train Set: 4.51% CER on Validation Set: 3.58%

OK Cancel

Hlavné zistenia

- ◆ Pri trénovaní malých modelov boli dosiahnuté veľmi dobré výsledky aj pri využití značne menšieho množstva tréningového materiálu
- ◆ Spoločné trénovanie paleograficky podobných rukopisov sa v mnohých smeroch javí efektívnejšie oproti ich samostatnému trénovaniu
- ◆ Využitie base modelu sa ukázalo ako jednoznačne výhodné pri vytváraní nových malých modelov, založených na jednom, ale aj viacerých rukopisoch



VISITATIO

Omnium Ecclesiarum, & Parochiarum
Archidiaconatus ZOLIENSIS,

Sanctimo D. N. ^{SUB} Benedicto XIV. P. M.

^{penes} Clementissimum.
MARIAE THERESIAE, DEI GR. R. H.
MANDATUM,

^{AC} Benignam
Celsissimi, ac Rdm. S. R. I. Principis, Dni. Domini Nicolai
e Comitibus CSAKY de Kereš, Perpetui Terrae Sc.
pusiensis; Miseratione Divina Ecclesiae Metropolitanae Stris.
Archi-Episcopi &c. &c.
Comissionem.

^{per}
Me Iosephum Comitem & BATTYAN, Perpetuum in Nemeth.
Ujvár, I. Collis Castri Ferrei Perpetuum Comitum, Protonotarium Apo.
Stolicum V. Capituli Ecclesiae Colles, S. Michaelis Archangeli de Castro
Ferro Praepositum, Excellmi. Illmi, ac Rdm. D. Francisci e Comitibus
Zichy de Vasonko Perpetui in Homonna ec. ec. Episcopi Laurinen. per
dictum Collum Castri in Spiritualibus Vicarium, & Caus. Audit. Stris.
Ecclesiae Metr. Strisoniensis Canonicum, Archid. Barschiensem,
peracta.

^{et}
Per Eund. quae V. Capituli Insis. Eccl. Coll. S. Martini Poson.
Praepositum, & Archid. necnon Eccl. Metr. Strisoni. Canonicum,
praesentata.

ANNÔ Domini,
M. DCC. LVI.

Automatický prepis
kanonickej vizitácie
Zvolenského
archidiakonátu
z roku 1754/1756

Kanonická vizitácie Zvolenského archidiakonátu Ostrihomského arcibiskupstva

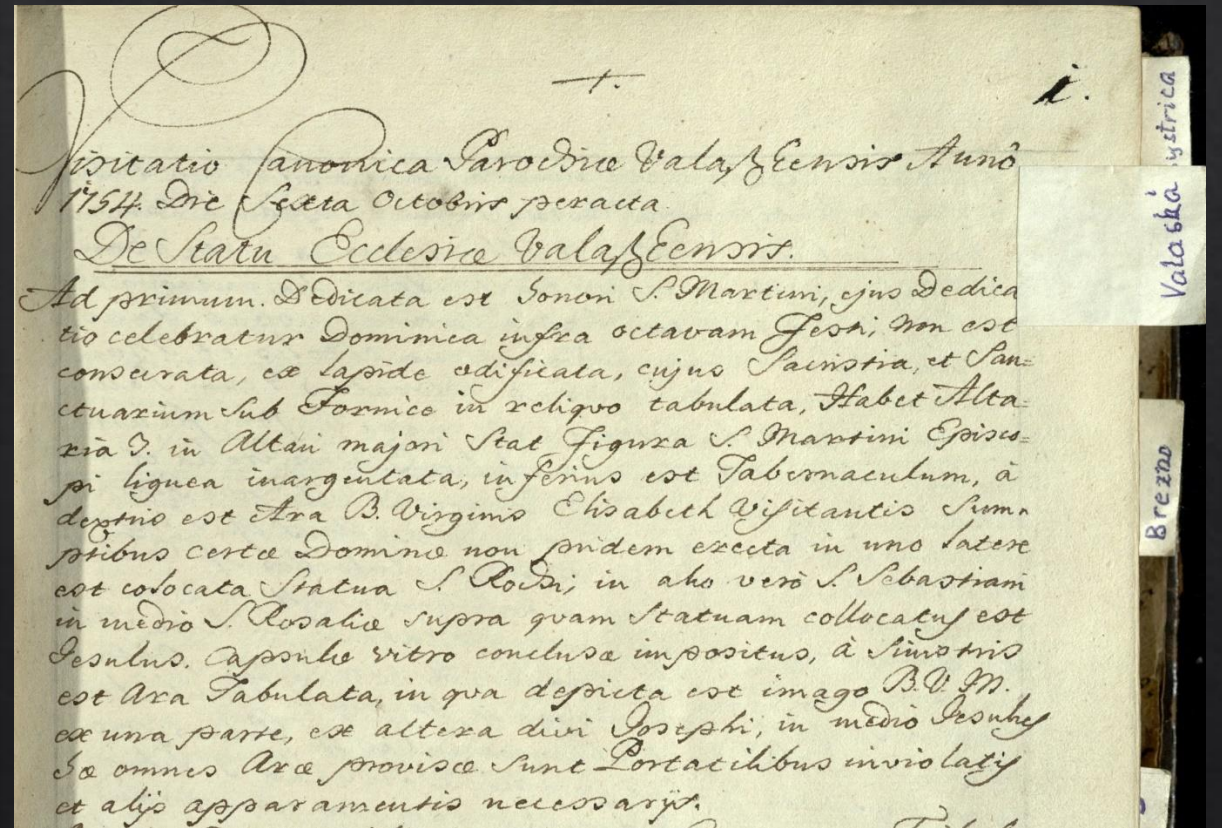
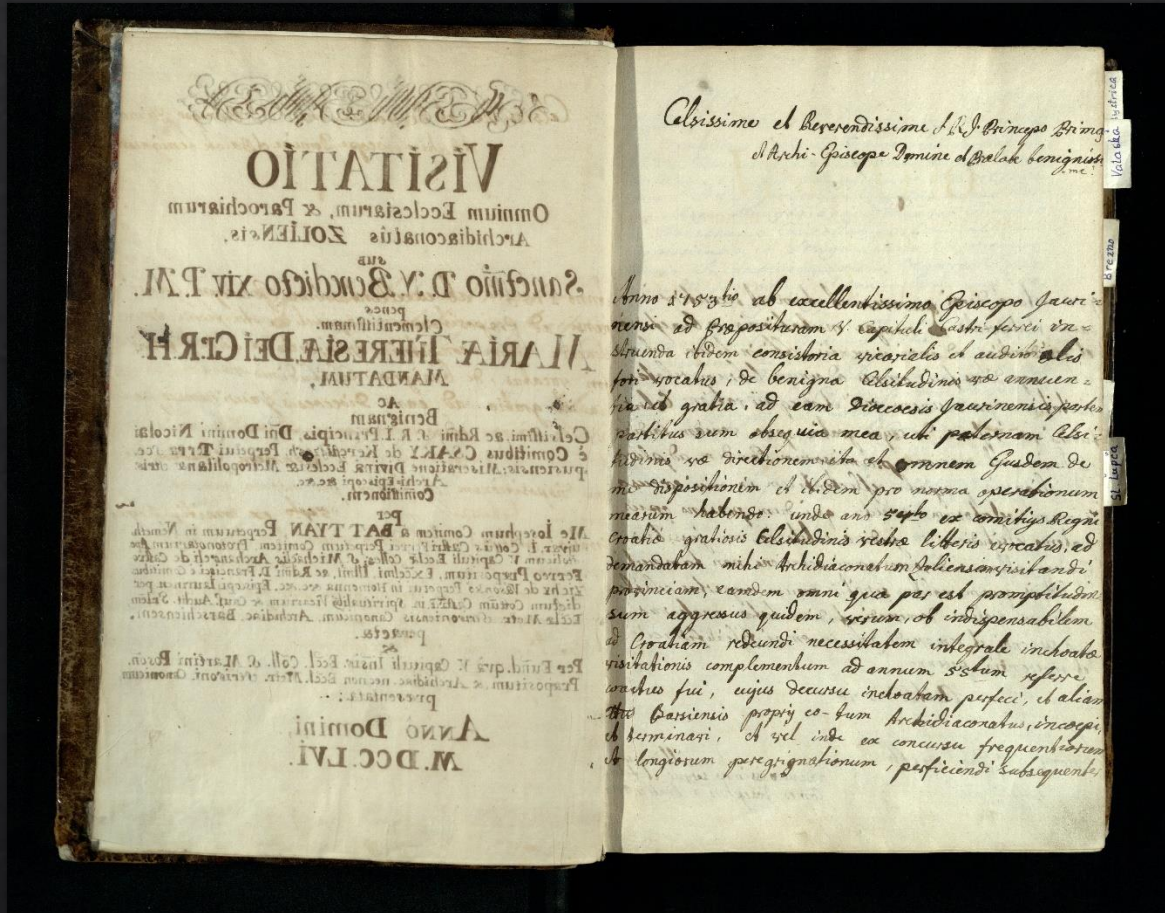
Čistopis vizitácie obsahuje cca 180 digitalizovaných dvojstrán textu.

Pozostáva z opisu vyše 30-tich farností a ich fílií, niekoľkých elenchov a indexu.

V texte sa vyskytuje niekoľko rukopisov, štyri rukopisy sú dominantné pri farnostiach:

1. Valaská, Lopej, H. Lehota, Predajná, Jasenie, Dubovica, Sv. Ondrej (súčasť Brusna), Ľubietová, Poniky, Dúbravica, S. Ľupča, Selce, Špania Dolina, Radvaň;
2. B. Bystrica;
3. Brezno;
4. Očová, Hrochoť, Detva, Slatinka, Zv. Slatina, Dobrá Niva, Sása, Krupina, Babiná, Zvolen, Budča, Bacúrov, Bačín, Sliač-Hájniky, Mičiná.

Text vizitácie obsahuje viacero rukopisov



Text vizitácie obsahuje viacero rukopisov

VISITATIO CANONICA.

Caempta Parochia Neosoliensis, & Filialium ejusdem
tredecim: Tajba, Kiecska, Godlavecz, S. Jacobus, Koshiviarzka,
Kordik, Kralik, Kaszova, Rudlova, Nemecz, Majer,
Kinczelova, Ulmanka, in Archi-Diaconatu Soliensi Districtu
Superioris, per me Josephum Gomilem à Battyan, C. M. S.
Canonicum, speciali Celsissimi S. R. I. Principis, Nicolai
Comilibus Clasy de Heresdszreggh / pleno Titulo: Archi-
Episcopi Krigonensis, Regni Hungariae Primatis De-
legatione instituta, Anno Millesimo Septingentesimo
Quingvagesimo Quarto.

Visitatio Ocsoviensis. De Statu Ecclesiae.

Ecclesia est dedicata honori Omnium Sanctorum, an sit Consecra-
ta? incertum est. Dedicatio tamen ejus celebratur Dominica infra
Octavam Festi. Tota lapidea, et formice Clara. Altaria habet tria. Me-
dium B. Virginis in coelis coronatae. Ad cornu Evangelii Obdormientis De-
atae Virginis; in cornu vero Epistolae S. Nicolai & Antonii Abbatis. O-
mnia tria Opus Antiquissimum, in quibus lapides portatiles integri, et
duo quidem cum sigillis externè apparentibus, in ternio non parat. ^{Ad}
Non scitur à quo fundata. Pro dote habet certam partem Territorij ex
cisam incipiendo à certo loco, infra Oppidum Zrudjensite dictorem
suo Territorium Zelnende ascendendo ad Territorium Hrocgotthyen-
se, locumq; ibidem nominatum po Trpkay, adeoq; hysbecke Dreziny Sol-
ne, Zorne, a Prostredne Srby, Bugarova az po Kshaciezu, ex altera vero
parte certi Agri Prostraya dicit; item Koshiczyne, quos via do Zelobuzeg
ducens ab Agris Dominati decime obnoxio dirigit. Intra limites
istos, quicquid indeminatur, decimat Ecclesia. Item habet Agros duos,
qui Ecclesia gemine ante paucos annos indeminabantur. ^{Ad}
_{finis}

Segmentácia a prepis vybraných pasáží vizitácie

Pre potreby vytvárania tréningových modelov vybrané tri rukopisy; z textu vizitácie boli prepísané dvojstrany:

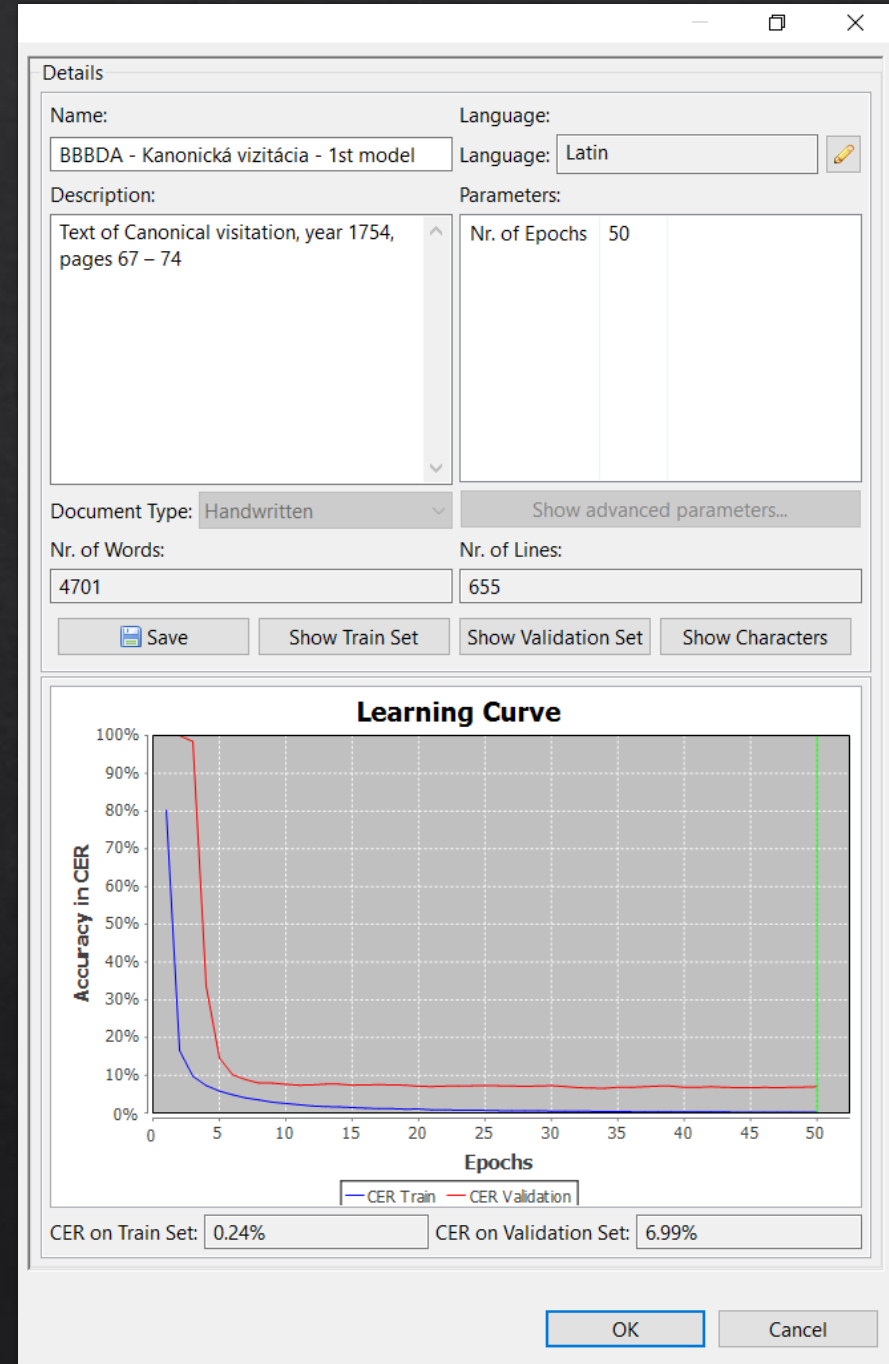
- ◇ 4 – 18 (farnosti Valaská a Lopej) → 2. model
- ◇ 67 – 75 (farnosť Banská Bystrica) → 1. model
- ◇ 101 – 113 (farnosti Očová a Hrochot') → 3. model (neskôr vylepšený)

Z prepisov jednotlivých rukopisov vytvorené tri malé tréningové modely automatickej transkripcie textu.

Výsledok 1. modelu

Rukopis vizitácie farnosti Banská Bystrica;
Celý rukopis má rozsah 16 dvojstrán (č. 67 -
82), pre potreby vytvorenia modelu
spracovaných 7 dvojstrán (4701 slov).

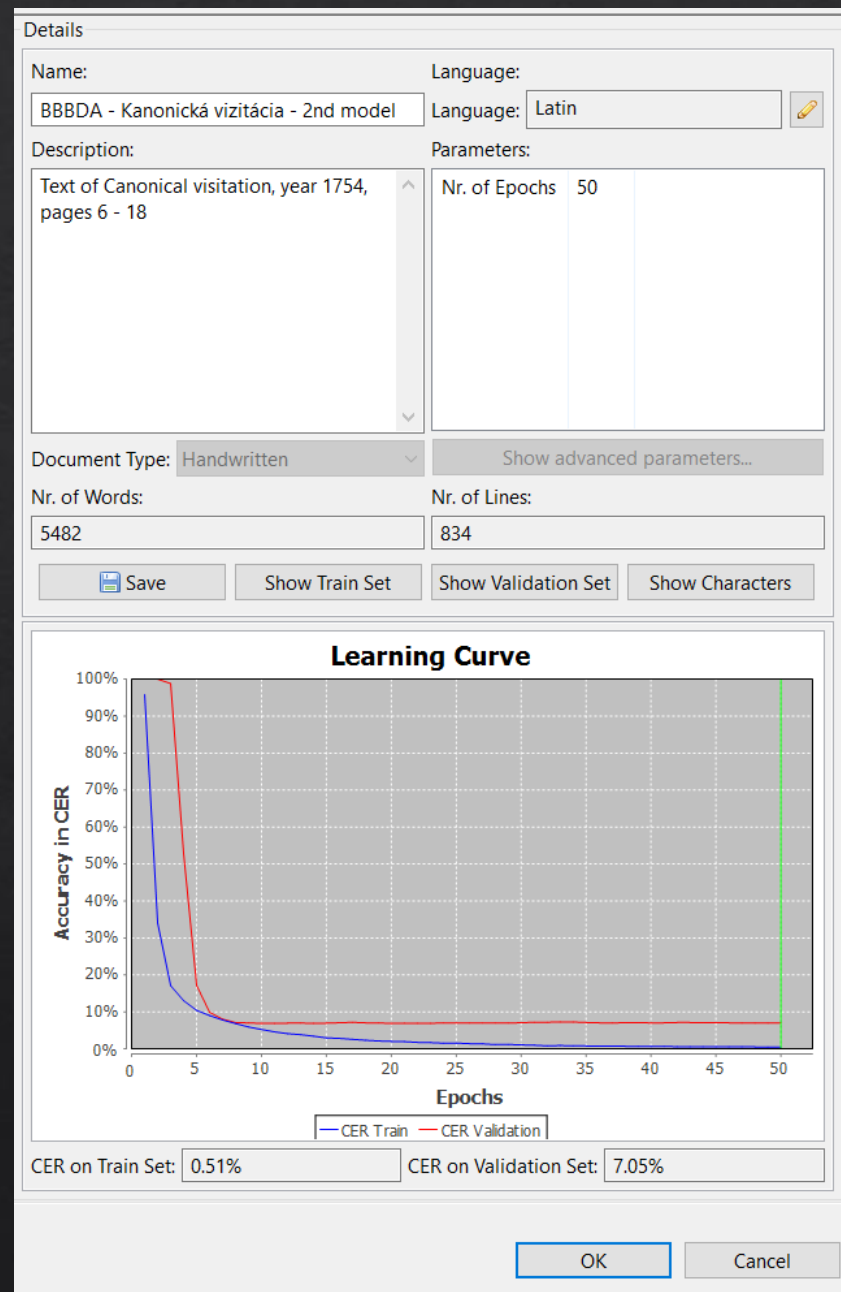
Hodnota CER vo validačnom súbore :
6,99 %



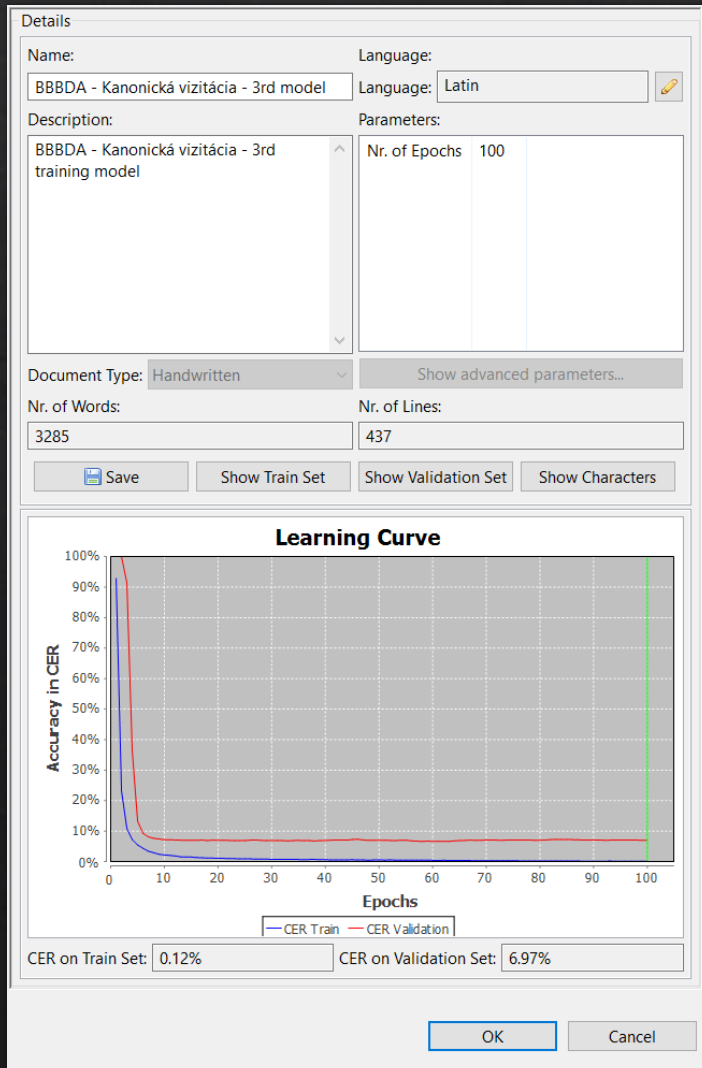
Výsledok 2. modelu

Rukopis vizitácie farností Valaská a Lopej; celý rukopis úvodnej časti vizitácie má rozsah 62 dvojstrán (č. 6 - 67), pre potreby vytvorenia modelu spracovaných 11 dvojstrán (5482 slov).

Hodnota CER vo validačnom súbore:
7,05 %

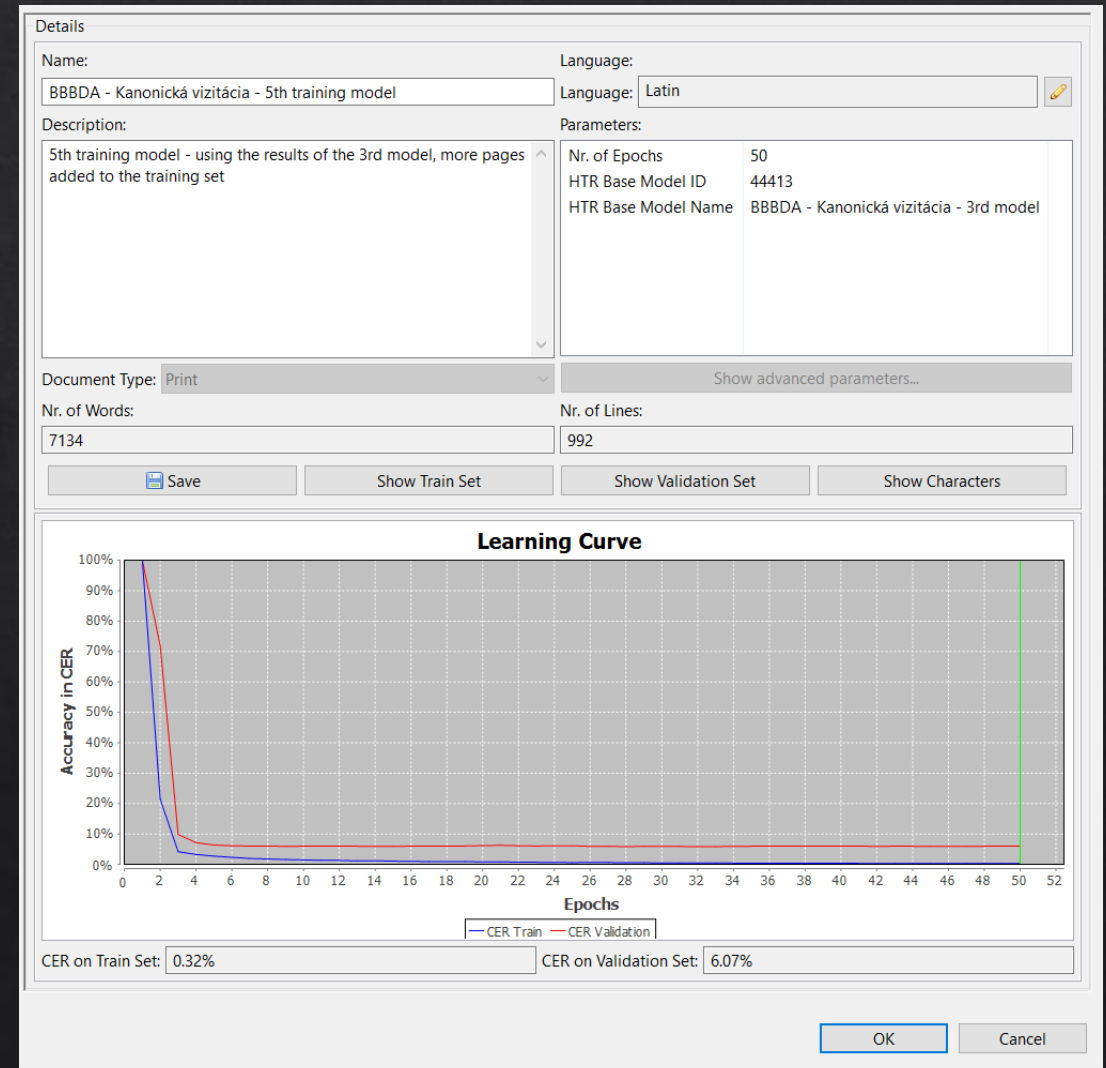


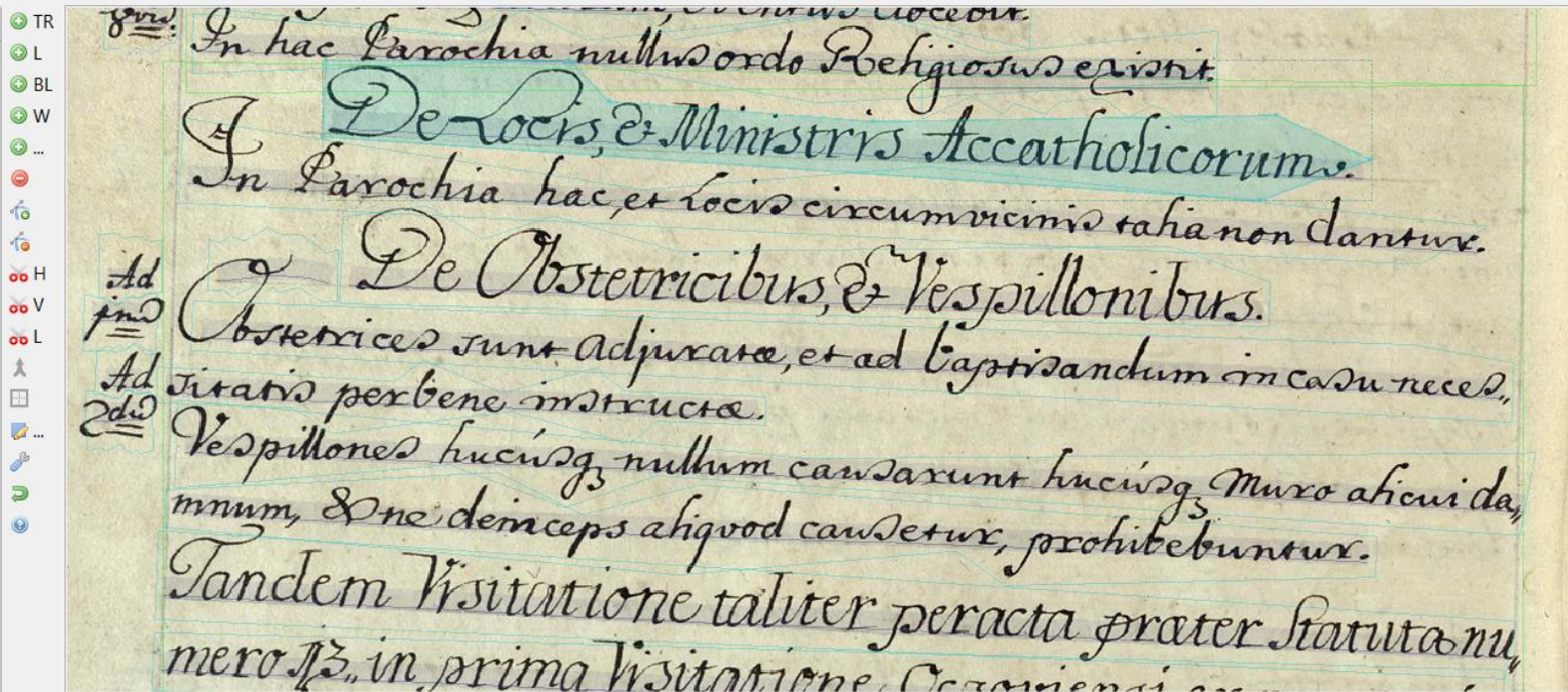
Výsledok 3. modelu a jeho vylepšenia



Rukopis vizitácie farností Očová a Hrochoť; pre potreby vytvorenia modelu spracovaných 7/13 dvojstrán (3285/7134 slov).

Hodnota CER vo validačnom súbore: 6,97 % a 6,07 %





1-20 In hac Parochia nullus ordo Religiosus existit.

2-1 De locis, et Ministris Accatholicorum.

2-2 n Parochia hac, et locis circumvicinis talia non dantur.

2-3 8

2-4 Ad

2-5 De Obstetricibus et Vespillonibus.

2-6 1mo

2-7 lbstetrices sunt adjuratae, et ad Captisandum in casu neces=

2-8 Ad

2-9 sitatis perbene instructae.