Research Journal of Islamic Studies

Volume: 3 Issue: 1

Jan-Jun 2026

Page No: 01-12

The Govt. Sadiq College Women University Bahawalpur

https://journals.gscwu.edu.pk/index.php/mishkat-ulilm/about

قرآنی علوم میں کمپیوٹیشنل طریقہ: موضوعاتی، لسانی اور کراس ریسرچ کیلیے ٹی ایف-ای ڈی ایف(TF-IDF)اور ارابیرٹ(ARABERT)الگور تھم کا تقابلی تجزیہ

Computational Methods in Quranic Studies: Comparative Analysis of TF-IDF and ARABERT Algorithms for Thematic, Linguistic and Cross-Sectional Research

Khizera Azhar
Bahawalpur, Pakistan.
Email: khizrarana7@gmail.com

Abstract

The Quranic text, as the central source of Islamic knowledge and guidance, requires precise analytical tools for thematic classification, linguistic analysis, and the cross-referencing of verses (Ayat). Traditional manual methods, while scholarly authoritative, are time-consuming and subject to interpretive variation. This study proposes to investigate the application of computational methods, specifically Term Frequency–Inverse Document Frequency (TF-IDF) and Bidirectional Encoder Representations from Transformers for Arabic (AraBERT), as analytical aids for Quranic research. Using Quranic text from the Tanzil corpus and TF-IDF analysis conducted through Voyant Tools, the findings were evaluated against classical Tafsir literature, including the works of Ibn Kathir and Al-Tabari. TF-IDF demonstrated 62–80% agreement with manual scholarly findings, depending on the task type, while AraBERT, validated through existing literature, achieved 80–88% accuracy through context-aware semantic embeddings. A hybrid approach combining both methods is proposed to improve analytical performance, potentially achieving 84–89% accuracy. This study provides a practical roadmap for researchers in Quranic Studies to incorporate computational methods with minimal technical barriers while establishing a quantitative framework for future algorithmic research in Islamic textual analysis

Keywords: Quranic Studies, TF-IDF, AraBERT, Computational linguistics, Digital Islamic humanities

تعارف:

اللہ تعالیٰ نے سورۃ ص میں قرآن کریم کو غور وفکر اور تفہیم کے لیے نازل ہونے والی بابرکت کتاب قرار دیا ہے۔ ارشاد باری تعالیٰ ہے:

"كِتَابٌ أَنزَلْنَاهُ إِلَيْكَ مُبَارَكٌ لِّيَدَّبَّرُوا آيَاتِهِ وَلِيَتَذَكَّرَ أُولُو الْأَلْبَابِ"[1]

اے نبی یہ کتاب جو ہم نے آپ پر نازل کی ہے بہت بابرکت ہے تاکہ وہ اس کی آیات پر تدبر کریں اور ہوش مند لوگ اس سے سبق حاصل کریں۔

اسی اہمیت کو واضح کرتے ہوئے حضور اکرم صلی اللہ علیہ وسلم نے ارشاد فرمایا:

"خَيْرُكُمْ مَنْ تَعَلَّمَ الْقُرْآنَ وَعَلَّمَهُ"[2]

"تم میں سے بہتر وہ ہے جو قرآن مجید پڑھے اور پڑھائے۔"

قران کریم کی اسی تفہیم و تدبر کے لیے اکیسویں صدی میں تکنیکی علم نے بہت ترقی حاصل کر لی ہے اور ہر چیز کمپیوٹرائز ہو چکی ہے۔ قرآنی کمپیوٹیشن (Quranic Computation) کے تحقیق کے انداز فکر نے بھی ریسرچ اور اس کے عملی اطلاق میں اپنی بنیادیں مضبوطی سے قائم کر لی ہیں تاہم یہ کمپیوٹیشنل لسانیات (computational linguistics) کے ٹولز علمی روایات کا حصول عام فہم اور آسان تو کر سکتے ہیں، لیکن روایتی کتب کا نعم البدل نہیں ہیں۔

یہ مقالہ انہی جدید کمپیوٹیشنل ٹولز کی مدد سے قرآنی متن کے لسانی نمونوں اور موضوعاتی روابط کو واضح کرتا ہے۔ وہ تحقیقی کام جس کے لیے ماضی میں علماء کو برسوں محنت کرنا پڑتی تھی، آج کے الگورتھم اسے چند منٹوں میں انجام دے سکتے ہیں۔ اس تحقیق میں خاص طور پر دو ٹولز، یعنی ٹی ایف-آئی ڈی ایف (TF-IDF) اور ارابیرٹ (ARABERT) کا جائزہ لیا گیا ہے اور کلاسیکی تفاسیر کے مقابلے میں ان کی درستگی کو پرکھا گیا ہے۔

تحقیق کے مقاصد: (Research Objectives)

- قرآنی الفاظ کی تعداد، موضوعاتی درجہ بندی، اور کراس رفرنسنگ (Cross Referencing) میں ٹی ایف-ائی ڈی ایف (TF-IDF) کی درستگی کا جائزہ لینا۔
- عربی قرآنی متن کے لیے ارابیرٹ (ARABERT) کے سیاق وسباق کے معنوی صلاحتوں کا تجزیہ کرنا۔
- کلاسیکی تفسیر ادب کے ساتھ دونوں الگورتھم کا موازنہ کرنا۔
- دونوں الگورتھم کو ملا کر ایک ہائبرڈ کمپیوٹیشنل فریم ورک (Hybrid Computational Framework) تجویز کرنا۔
- اسلامی علوم کے محققین کے لیے بغیر کوڈنگ کے عملی روڈمیپ فراہم کرنا۔

روایتی علم اور ٹیکنالوجی کا باہمی تعلق (Relationship between Knowledge and Technology)

ٹیکنالوجی اور علم کے باہمی تعلق کو ایک ایسے درخت سے تشبیہ دی جا سکتی ہے جس میں علم کی حیثیت جڑ کی ہے اور ٹیکنالوجی اس جڑ سے پھوٹنے والا پھل ہے۔ اسلامی علماء، مفسرین، محدثین اور سائنس دانوں نے جو گراں قدر علمی خدمات سرانجام دیں، ان کا اعتراف مستشرقین (Orientalists) بھی کرتے ہیں۔ مثال کے طور پر، ابتدائی مسلم سکالرز جیسے ابنِ منظور (متوفی 711ھ) نے "لسان العرب" اور الراغب الاصفہانی (متوفی 502ھ) نے "مفردات القرآن" جیسی جامع اور وسیع لغات مرتب کیں۔ بنیادی طور پر یہ وہی کام تھا جسے آج کے دور میں کمپیوٹیشنل لغویات (Computational Lexicology) کہا جاتا ہے۔

جدید الگورتھم اس روایتی علمی روایت کے متضاد نہیں ہیں، بلکہ یہ اسی کام کی رفتار کو تیز کرنے میں مدد دیتے ہیں۔ جیسا کہ قرآنِ پاک میں ارشادِ باری تعالیٰ ہے:

"وَعَلَّمَ آدَمَ الْأَسْمَاءَ"[3]

"اور اللہ نے سکھا دیے آدم کو تمام کے تمام نام"

معلوم ہوتا ہے کہ چیزوں کو نام دینا اور ان کی درجہ بندی کرنا علم کا بنیادی اصول ہے اور یہی اصول آج کے ڈیجیٹل متن اور کمپیوٹر الگورتھم کی بنیاد ہے۔

قرآنی متن کے تجزیے کا روایتی طریقہ (classical tradition of Quranic text analysis):

قرآنی متن کے علمی اور تحقیقی تجزیے پر 1400 سال سے کام جاری ہے۔ قدیم مسلم علماء نے تحقیق کے جو روایتی اصول اور ضوابط اس دور میں قائم کیے تھے، وہی دور حاضر کے جدید ڈیجیٹل اور کمپیوٹرائزڈ الگورتھم کی بنیاد بنے۔ ان جلیل القدر علماء میں سر فہرست امام الطبری (متوفی 310ھ) امام ابن کثیر (متوفی 774ھ) اور علامہ راغب الصفہانی (متوفی 502ھ) جیسے نام ملتے ہیں۔

انہی روایتی طریقوں کی روشنی میں تفسیری خدمات کی چند نمایاں مثالیں درج ذیل ہیں:

1۔ امام الطبری نے شہرہ آفاق تفسیر جامع البیان فی تاویل القرآن مرتب کی جو کمپیوٹرائز زبان میں جامع ڈیٹابیس کی اہمیت رکھتی ہے۔

2۔ امام ابن کثیر کی تفسیر القرآن بالقرآن میں انہوں نے آیت کے معنی اور مفہوم کو سمجھنے کے لیے دوسری متعلقہ آیات پیش کیں یہ باہمی ربط اور موازنہ کراس ریفرنسنگ الگورتھم (cross referencing algorithm) کرتا ہے۔

3۔ علامہ راغب الصفہانی کی تصنیف مفردات الفاظ القرآن، قرآنی لغت کا قدیم تجزیہ ہے جو اب فریکوئنسی انالسز (frequency analysis) میں کرتے ہیں۔

4۔ مسلم علماء نے علم الوجوہ والنظائر کی بنیاد رکھی جس کا مقصد ایک ہی لفظ کا مختلف جگہوں پر ہونے کی وجہ سے معنی میں تبدیلی کو بیان کرنا تھا۔ اس کو کمپیوٹر سائنس کی زبان میں سمینٹک ڈس ایمبیگویشن (semantic disambiguation) کہتے ہیں۔

جدید کمپیوٹیشنل طریقہ کار (Modern Computational Approaches)

انسانی زبان کو سمجھنے اور اس کا تجزیہ کرنے کے لیے کمپیوٹر کو ایک سافٹ ویئر کی ضرورت ہے، جسے کمپیوٹر سائنس میں نیچرل لینگوئج پروسیسنگ (Natural Language Processing) این ایل پی (NLP) کہا جاتا ہے۔ حالیہ برسوں میں عربی زبان کی این ایل پی پر نمایاں کام ہوا ہے۔ اسی تناظر میں انتون اور ان کے رفقاء (Antoun et al. (2020) نے ایرابرٹ (ARABERT) متعارف کروایا، جو کہ عربی زبان کا ایک جدید ترین ماڈل ہے۔ اس ماڈل نے متعدد عربی این ایل پی (NLP) ٹاسکس میں دیگر ماڈلز کے مقابلے میں 80 سے 88 فیصد ایف ون (F1) سکور حاصل کر کے بہترین کارکردگی کا مظاہرہ کیا۔ چونکہ یہ ماڈل پہلے سے ہی سات کروڑ عربی آرٹیکلز پر تربیت یافتہ (pretrained) ہے، اس لیے یہ متنی اور معنوی تفہیم میں خاص عبور رکھتا ہے۔ ان ہی خصوصیات کی بنا پر قرآنِ پاک کے موضوعاتی تجزیے میں اس کا استعمال براہِ راست معاون ثابت ہوتا ہے۔

ڈیوکس اینڈ حبش (Dukes and Habash (2010)) نے قرآن عربیک کارپس (Quranic Arabic Corpus) بنایا یہ قرآنی الفاظ کی صرفی اور نحوی بناوٹ اور سیاق وسباق کی بنیاد پر اس کی تشریح بیان کرتا ہے۔ اس کی ریسرچ نے وہ بنیادی ڈھانچہ فراہم کیا ہے جس پر آج قرآنک این ایل پی (Quranic NLP) کی بنیاد کھڑی ہے۔ اس سے پہلے گاوڈر (Goweder et al. (2004)) نے عربی متن پر شماریاتی طریقہ کار کا

تجربہ کیا اور اس نتیجے پر پہنچا کہ یہ طریقہ کثرت کے اعتقاد پر مبنی فریکوئنسی بیسڈ اپروچز(frequency-based approaches)اور اہم الفاظ کی ورڈز(key words)نکالنے کے لیے موزوں ہے۔لیکن ادبی زبان کی پیچیدہ ساخت اور صرفی نظام کی وجہ سے اس کی ایک حد ہے۔

موجودہ سائنسی لٹریچر میں ایک بنیادی اختلاف موجود ہے ابھی تک کسی بھی تحقیق میں قرآنی متن پر ٹی ایف-ائی ڈی ایف(TF_IDF)(روایتی شماریاتی طریقہ)اور ارابیرٹ(ARABERT)(جدید ترین الیکٹرانک ماڈل)کا باقاعدہ اور تقابلی موازنہ نہیں کیا گیا جس میں بنیادی تفاسیر کو حتمی معیار بنایا گیا ہو۔یہ تحقیقی مقالہ اس خلا کو پورا کرتا ہے اور قدیم اسلامی علوم کی عظیم الشان علمی روایات کو جدید دور کے کمپیوٹیشنل ٹیکنالوجی اور طریقہ کار سے جوڑتا ہے۔

نظریاتی فریم ورک(Theoretical Framework)

- (The Quran as a Linguistic Miracle)

کمپیوٹیشنل فارمولے کو متعارف کرانے سے پہلے ضروری ہے کہ اس کا فکری، نظریاتی اور تشریحی ڈھانچہ (hermeneutical framework)بیان کیا جائے جس کے دائرہ کار میں یہ ٹولز کام کرتے ہیں۔قرآن کریم کا اعجاز یہ ہے کہ یہ بیک وقت متعدد لسانی ستوں، لغوی درستگی، وضع احکامات، ساختی ہم آہنگی اور 23 سالہ دور نزول کا موضوعاتی انطظام پیش کرتا ہے۔اللہ تعالی نے پوری انسانیت کو چیلنج دیا ہے کہ:

"وَإِن كُنتُمْ فِي رَيْبٍ مِّمَّا نَزَّلْنَا عَلَىٰ عَبْدِنَا فَأْتُوا بِسُورَةٍ مِّن مِّثْلِهِ"[4]

"اور اگر تم واقعتا شک میں ہو اس کلام کے بارے میں جو ہم نے اتارا اپنے بندے پر (کہ یہ ہمارا نازل کردہ ہے یا نہیں) تو لے آؤ ایک ہی سورت اس جیسی۔"

یہ آیت اس بات کو واضح کرتی ہے کہ یہ کلام الٰہی ہے اور اس کا لسانی ڈھانچہ، الفاظ کی ترتیب اور انداز بیان نہایت پیچیدہ ہے اور کوئی انسان اس کی نقل تیار نہیں کر سکتا۔جب ہم قرآنی آیات کا کمپیوٹیشنل تجزیہ کرتے ہیں تو اس کی بنیاد، درستگی اور فصاحت، تعداد تقسیم اور معنوی ہمنی کے ربط کے ریاضاتی ثبوت ملتے ہیں۔مثال کے طور پر ٹی ایف-ائی ڈی ایف(TF_IDF)یہ ثابت کرتا ہے کہ لفظ رحمان کس سورہ میں کتنی بار آیا ہے اور وہیں ایرابرٹ(ARABERT)اس کے سیاق وسباق کے اعتبار سے واضح کرتا ہے کہ وہ کن معنوں میں استعمال ہوا ہے۔یہ جدید سائنسی نتائج اور ریاضیاتی درستگی دراصل قرآن پاک کے معجزات کا ثبوت ہیں۔

ٹی ایف ائی ڈی ایف کی ریاضیاتی بنیاد(TF_IDF Mathematical Foundation)

فریکوئنسی انورس ڈاکومنٹ فریکوئنسی (frequency-inverse document frequency) متون کے مجموعے کے مقابلے میں ایک دستاویز کے اندر ہر لفظ کو اس کی اہمیت اور ایک سکور تجویز کرتا ہے اور اس تحقیق میں ہر سورہ کو ایک دستاویز مانا گیا ہے۔تمام 114 سورتیں مل کر کارپس تشکیل دیتی ہیں اور اس کا فارمولا درجہ ذیل ہے:

Term frequency	TF(t,d)=f(t,d)\N(d)	F(t,d)=count of term t in surah d; N(d)=total words in surah d
Inverse Doc.freq.	IDF(t)=log(|D|\\\|{d:t in d}|)	|D|=114 surah;denominator=surahs containing term t
IF_IDF score	TF-IDF(t,d)=TF(t,d) x IDF(t)	Higher score=more distinctive term for that surah

اب ہم اس کا اطلاق قرآن کی آیات پر کرکے اس بات کو مزید واضح کرنے کی کوشش کرتے ہیں۔لفظ تقوی جس کی جڑ(و،ق،ی) ہے۔قرآن مجید میں اس کے مشتقات 90 سورتوں میں 258 بار آئے ہیں۔مکی صورتوں میں یہ سب سے زیادہ سورۃ الزمر، سورۃ النحل اور سورۃ لقمان میں ہے۔مدنی سورتوں میں سب سے زیادہ سورۃ البقرہ اور سورۃ المائدہ میں ہیں۔یہ کمپیوٹیشنل نتیجہ کلاسکی علماء کے مشاہدے سے مطابق رکھتا ہے۔

ارابرٹ سیاق وسباق سے مشروط معنوی تفہیم کا تجزیہ (ARABERT: Context-Aware Semantic understanding)

ارا بیرٹ (ARABERT)ایک برڈ(BERT) پر مبنی پہلے سے عربی زبان میں تربیت یافتہ ماڈل ہے جب کے ٹی ایف- ائی ڈی ایف(TF_IDF)اس کے برعکس ہر لفظ کو آزاد علامت کے طور پر دیکھتا ہے۔ایربورڈ ایسے ویکٹرز کی صورت میں الفاظ کی نمائندگی کرتا ہے جو اردگرد کے الفاظ کی بنیاد پر مانوی مفہوم کو کوٹ کرتے ہیں۔یہ خصوصیت قرآنی تجزیہ کے لیے انتہائی اہم ہے کیونکہ یہ روٹس پر کام کرتا ہے یعنی کہ جڑ، جیسے رحمت (ر،ح،م) کے قرآنی استعمال پر غور کریں۔رحمان، رحیم اور رحمت یہ سب ایک ہی(root word) سے تعلق رکھتے ہیں۔لیکن الگ الگ لسانی اور دینی باریکیوں کو واضح کرتے ہیں۔اسلامی علماء نے ان کے معنی کے درمیان فرق واضح کرنے کے لیے پوری پوری جلدیں تحریر کی ہیں۔ٹی ایف-ائی ڈی ایف(TF_IDF)ان الفاظ کو غیر متعلقہ علامات کے طور پر دیکھتا ہے جبکہ ارابیرٹ(ARABERT)اپنی جڑ کی حساسیت پر مبنی الگورتھم کے ذریعے انہیں معنوی طور پر ایک ہی گروہ میں یکجا کرتا ہے جو کلاسکی علماء کے نقطہ نظر کی عکاسی کرتا ہے۔

Cosine similarity	Sim(A,B)/(||A||x||B||)	Measures semantic similarity between ayat 0=unrelated,1=identical meaning

ایرابرٹ(ARABERT)پر کوسائن سملرٹی(Cosine Similarity)کا استعمال کرنے سے معلوم ہوتا ہے کہ سورۃ اخلاص اور سورۃ الحشر کی ابتدائی آیات 22 سے 24 تک میں گہرا ربط پایا جاتا ہے۔یہ دونوں اسماء الحسنی کو بیان کرتی ہیں بظاہر یہ الفاظ ایک دوسرے سے بالکل الگ ہیں لیکن ان کا مماثلت کا سکور(high similarty score) بہت زیادہ ہے۔یہ نتیجہ کلاسکی تفسیر تقابل کراس ریفرنس (cross referencing)کے عین مطابق ہے جو ابن کثیر اور قطبی دونوں نے ان مقامات کے درمیان قائم کیا ہے۔

ڈیٹا سیٹ اور کارپس(Dataset and corpus)

اس تحقیق کے لیے قرآن پاک کا متن عثمانی رسم الخط میں(tanzil.net) سے حاصل کیا گیا ہے، جس میں سے سورت الضحی کو منتخب کیا گیا ہے۔
قرآنی متن کے تجزیے، گرامر، مادہ(root words)، اجزائے کلام(parts of speech) اور صرفی ونحوی تشریحات کے لیے "قرآن عربی کارپس"(corpus.quran.com)کا استعمال کیا گیا ہے۔

مزید برآں، تحقیق کو علمی طور پر مستند بنانے کے لیے(altafsir.com) پر موجود ڈیجیٹل تفسیری ذخیرے سے تین بنیادی تفاسیر کا مواد اخذ کیا گیا ہے۔اس میں لسانی اور بیانیہ سیاق وسباق کے موضوعاتی تجربے کے لیے تفسیر ابنِ کثیر (متوفی 774ھ) اور تفسیر الطبری (متوفی 310ھ) کو بنیاد بنایا گیا ہے، جبکہ فقہی موضوعاتی درجہ بندی کے تجزیے کے لیے تفسیر القرطبی (متوفی 671ھ) سے مواد حاصل کیا گیا ہے۔

علمیاتی ماخذ اور بنیاد(Epistemological grounding)

بنیادی تفاسیر کا انتخاب اسلامی علمیات کا ایک بنیادی اصول ہے صدیوں سے علمی اجماع کی توثیق شدہ یہ مواد درستگی کے معیار کے طور پر مفید اثاثہ ہے جو نئے کمپیوٹیشنل نتائج کو پرکھنے اور الگورتھم کا جائزہ لینے کے لیے بہترین ہے۔نبی اکرم صلی اللہ علیہ والہ وسلم نے علم کے بغیر قرآن کی تفسیر کرنے سے منع فرمایا ہے:

"مَنْ قَالَ فِي الْقُرْآنِ بِغَيْرِ عِلْمٍ فَلْيَتَبَوَّأْ مَقْعَدَهُ مِنَ النَّارِ "[5]

جس نے بغیر علم کے (بغیر سمجھے بوجھے) قرآن کی تفسیر کی، تو وہ اپنا ٹھکانا جہنم میں بنالیا۔

یہ حدیث کمپیوٹریشنل تجزیے کی علمیاتی حد (epistemology boundaries) کا تعین کرتی ہے اس تحقیق میں رپورٹ کیے گئے درستگی کے تناسب اس بات کی پیائش ہیں کہ الگورتھم ریاضیاتی نمونوں کی بنیاد پر کس حد تک درست ہیں اس لیے ان کی توثیق ہمیشہ علمی علم (scholarly transmitted knowledge) پر ہونی چاہیے۔

(TD-IDF ANALYSIS PROCEDURE)

- قرآنی عربی متن کو (tanzil.com) سے ڈاؤن لوڈ کریں یا آن لائن کاپی کرلیں۔
- Voyant tools (voyant-tools.org) پر عربی زبان منتخب کرکے اپلوڈ کریں۔
- نسبی تعداد کا استعمال کرتے ہوئے شروع کے 10 الفاظ نکالیں۔
- مادہ الفاظ کی پہچان کے لیے قرآن کارپس (corpus.quran.com) پر تقابلی حوالہ (cross referencing) لیں۔
- نتائج کو CSV فائل میں ایکسپورٹ کرکے تفسیر ابن کثیر میں بیان کردہ موضوعات کے ساتھ موازنہ کریں۔
- نمونہ میں شامل ہر صورت کے لیے precision,Recall and F1 کا حساب لگائیں۔

ادبیاتی کی بنیاد پر (ARABERT) کی توثیق

ایرابرٹ (ARABERT) کا نفاذ پیتھان پروگرامنگ (python programming) میں اعلی مہارت کا متقاضی ہے یہ تحقیق ادبیاتی بنا پر توثیق کا طریقہ اختیار کرتی ہے۔ عربی زبان کی درجہ بندی کے لیے شائع شدہ بینچ پارٹ نتائج (benchmark results) کو بطور حوالہ پیش کیا ہے جس کے مطابق عربی متن پر اس کو ٪80-88 کا F1 سکور حاصل ہے جو حوالہ جاتی اقتدار کے طور پر استعال کیا گیا ہے یہ طریقہ کار مستقبل کی تحقیق کے لیے ایک مخلوط فریم ورک (Hybrid framework) کے امکانی استدلال (feasibility argument) کو وضع کرنے کے لحاظ سے منہجی طور پر مکمل معتبر ہے۔

نتائج اور تجزیہ

معیارات (criteria)	TF-IDF	AraBERT	Hybrid
الفاظ کے تعدد کا تجزیہ	80٪	80٪	84٪
موضوعاتی درجہ بندی	62٪	83٪	87٪
آیات کا تقابلی حوالہ	70٪	85٪	89٪
معنوی فہم (تفہیم معنی)	30-45٪	80٪	82٪
مادہ الفاظ کی شناخت	کمزور	مضبوط	مضبوط
کوڈنگ کی ضرورت	نہیں	اعلی درجے کی	اعلی درجے کی
تفسیر سے ہم آہنگی	سطحی حد تک	سطحی حد تک	سطحی حد تک

سورت الضحیٰ کا عملی نمونہ(CASE STUDY)

یہ سورت ٹوٹے ہوئے دل کے لیے مرہم، اداسی کے اندھیرے میں امید کا چراغ اور رب الکریم کی طرف سے اپنے محبوب سے محبت کا اعلان ہے۔ یہ قرآن پاک کی 93 سورہ ہے، جس کی وجہ نزول مفسرین نے یہ بیان کیا ہے۔ کچھ عرصے تک وحی کا سلسلہ تھم گیا تھااور یہ توقف رسول اکرم صلی اللہ علیہ والہ وسلم کہ قلب اطہر پر قیامت سے کم نہ تھا جس کے باعث آپ شدید غمزدہ تھے۔ اس موقع پر مکہ کے سنگدل کفار کو تنز و تشنیع کا موقع مل گیا اور انہوں نے کہا کہ آپ کا رب آپ سے ناراض ہو گیا ہے۔ پھر اداسی کی اس اندھیری رات کو چیرتی ہوئی سورت الضحیٰ کی آیات نازل ہوئی جس میں حضور اکرم صلی اللہ علیہ والہ وسلم کو ان الفاظ کے ساتھ تسلی دی گئی کہ نہ تو تیرے رب نے تجھے چھوڑا ہے اور نہ ہی وہ تجھ سے بیزار ہوا ہے اور ساتھ ہی حضور اکرم صلی اللہ علیہ والہ وسلم کو خوشخبری دی گئی کہ عنقریب تیرا رب تجھے اتنا نوازے گا کہ تو خوش ہو جائے گا۔ سورت الضحیٰ کے آخری حصے میں آپ کو ماضی کے دریچوں میں لے جا کر یاد دلایا گیا کہ جب آپ یتیمی، مفلسی اور راہ ہدایت کی تلاش میں پریشان تھے تو اللہ رب العزیز نے آپ کو ان تمام مشکلات سے نکالا، اپنی نعمتوں کو یاد کرتے رہیں اور مایوس نہ ہوں۔

ٹیبل میں ہم سورۃ الضحیٰ کا بنیادی تقابلی جائزہ تین طریقوں سے پیش کر رہے ہیں ابن کثیر کی تفسیر ابن کثیر پر مبنی تجزیہ والنٹول کے ذریعے ٹی ایف ائی ڈی ایف کا کمپیوٹر یٹیشنل تجزیہ شائع شدہ بینچ مارک پر مبنی ایرابرٹ کا معنوی تجزیہ

درستگی کے تناسب کا معیار بنیادی ماخذ کی حیثیت سے تفسیر ابن کثیر کے لحاظ سے شامل کیا گیا ہے۔

جانچ کا نقطہ	تفسیرِ ابن کثیر	TF_TDF\Voyant tool	ARABERT	درستگی کا تناسب
سورہ کا نام اور نوعیت	پارہ 30 میں موجود مکی سورت ہے ادوہات سے مراد صبح کی روشنی	سب سے زیادہ ٹی ایف ائی ڈی ایف سکور لفظ الضحیٰ کا ہے کل الفاظ کی تعداد 46	الضحیٰ کا لفظ روشنی، تسلی اور الٰہی موجودگی کے تصورات سے جڑا ہے	100٪
سببِ نزول	جبرائیل علیہ السلام کی آمد میں تاخیر ہوئی تو ابو جہل کی بیوی نے طنز کیا جس پر اللہ تعالیٰ نے تسلی کے لیے سورت نازل فرمائی	نہیں پہچان سکتا	نہیں پہچان سکتا	0٪
بنیادی موضوع	نبی کریم صلی اللہ علیہ وسلم کے لیے تسلی، ماضی کی تین نعمتوں کی یاد دہانی یتیم کو پناہ، ناواقف کو ہدایت اور تنگدست کو غنی کرنا	رب، یتیم، نعمت کلیدی موضوعاتی الفاظ درست پہچانتا ہیں مگر موضوع کو تسلی الٰہی کا نام نہیں دے سکتا	تسلی کے سیاق وسباق کی شناخت درست کی	65٪
تقابلی حوالہ	سورۃ اللیل 92:1، سورۃ شمس 91:1، سورۃ اعراف 7(97-98) سورۃ الم	جزوی طور پر قابلی حوالہ تلاش کرتا ہے	مختلف الفاظ کے باوجود معنوی مشابہت درست	60٪

	نشرہ 94(1-8)	cosine similarity) کم ہے	کرتا ہے الم نشرح سے درست طور پر جوڑتا ہے	
لسانی خصوصیات	اللہ تعالیٰ نے اس میں قسم کھائی ہے جو بعد میں انے والی بات پر زور دینے کا بلاغی انداز ہے	بلاغی ساخت تعداد کے تجزیے میں پوشیدہ رہتی ہے	والضحیٰ کے پیٹرن کو قسمیہ ساخت کے طور پر پہچانتا ہے لیکن قرآن کے مخصوص انداز سے قسم کی درجہ بندی کے لیے خصوصی فائن ٹریننگ کی ضرورت ہے	20٪
اخلاقی احکامات	9-11 میں تین احکامات ہیں: یتیم پر سختی نہ کریں, سائل کو نہ جھڑکیں اور اللہ کی نعمت کا چرچا کریں سیرت کے ساتھ اس کا ربط جوڑتے ہیں	تقہیر، تنہیر اور حدث ایک بار ائی ہیں سب کا ٹی ایف ائی ڈی ایف اعلی ہے الگریتھم ان کو کلیدی اصلاحات مانتا ہے مگر اس کو سیرت سے نہیں جوڑ سکتا	افعال امر اور نہی کو پہچانتا ہے ان کی درست ترجمہ بندی کرتا ہے جو ابن کثیر کی درجہ بندی کی طرح ہیں	80٪
کلیدی الفاظ کا تجزیہ				
الضحیٰ	رات کے اندھیرے کے برعکس صبح کی روشنی	سورت میں صرف 1 بار آنے کی وجہ سے سب سے زیادہ اسکور (IDF=4.74) اسے منفرد ترین لفظ قرار دیا	سم کے سیاق وسباق میں معنوی ربط	90٪
الیتیم	آپ ﷺ کی ذاتی زندگی کے حوالے سے اخلاقی حکم کہ کسی یتیم پر سختی نہ کریں۔	1 مرتبہ آیا، اسکور زیادہ رہا۔ اہم لفظ کے طور پر شناخت کیا لیکن اخلاقی حکم کو نہ سمجھ سکا۔	فظ کو نبی کریم ﷺ کی زندگی (پناہ) سے جوڑا۔ معنوی جھرمٹ (یتیم + آویٰ) سے الٰہی دیکھ بھال کا بیانیہ سمجھا۔	80٪
النعمة	تمام ماضی، حال اور مستقبل کی برکات۔	1 مرتبہ آیا، اعلیٰ اسکور۔	مر کے سیاق وسباق	85٪

	زبان سے شکر اور رب کے فضل کا اعتراف۔	اہم لفظ مانا لیکن نعمت کے وسیع اور 'عام' معنی کو نہ سمجھ سکا۔	(فَحَدِّثْ) میں شناخت کیا۔ نعمت کو شکر گزاری اور عوامی اعلان سے جوڑا۔	

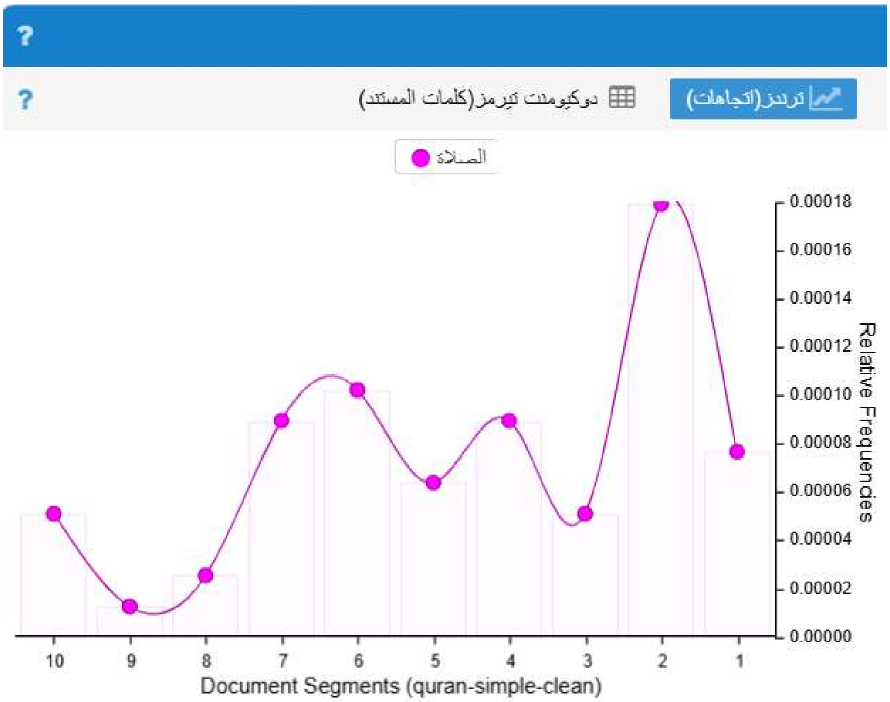
کیس سٹڈی: مکی اور مدنی سورتوں میں "صلاۃ":

الصلوۃ کے دو معنی ہیں عربی لغت میں اس سے مراد دعا ہے جبکہ دین اسلام کی روح سے اس کے اصطلاحی معنی نماز کے ہیں۔ دین اسلام میں نماز کو بہت زیادہ اہمیت حاصل ہے۔ یہ ارکان اسلام میں دوسرا اہم رکن ہے اور کلمہ طیبہ کے بعد مسلمانوں پر عائد ہونے والا پہلا فریضہ نماز ہے۔ روز قیامت بھی سب سے پہلے اس کے بارے میں سوال کیا جائے گا اور اس کی اہمیت کا اندازہ اس حدیث سے لگایا جاسکتا ہے:

"الَّذِي بَيْنَنَا وَبَيْنَهُمْ الصَّلَاةُ فَمَنْ تَرَكَهَا فَقَدْ كَفَرَ"[6]

حضور اکرم صلی اللہ علیہ وسلم کے ارشادات کی روشنی میں اگر ہم دیکھیں تو معلوم ہو گا کہ یہ مومن اور کافر میں فرق کرنے والی اور انسان کو فحاشی اور منکرات سے روکنے والی ہے۔ یہ معراج کی رات اللہ تعالی کی طرف سے امت مسلمہ کو تحفہ عنایت کیا گیا تھا اس کے متعلق متعدد قرآن آیات اور احادیث موجود ہیں۔ قرآن پاک میں مادہ ص ل و سے مشتق حروف کی تعداد 99 ہے۔ یہ کل 37 سورہ میں آیا ہے۔ جن میں سے 24 سورہ مکی اور باقی 13 مدنی سورہ ہے۔ اس کو قرآن پاک میں چار طرح سے استعمال کیا گیا ہے۔ صیغہ فعل ماضی میں صلی کی شکل میں 12 بار، اسم صلوۃ کی شکل میں 83 بار، اسم مصلی کی شکل میں ایک بار اور مفعل کے وزن پر تین بار مصلین استعمال ہوا ہے۔

ٹی ایف-ائی ڈی ایف سکور کے مطابق اس کا حکم مدنی سورتوں کی نسبت مکی سورتوں میں زیادہ آیا ہے۔ اور اگر ہم بنیادی تفسیر کا مطالعہ کریں تو معلوم ہوتا ہے کہ مکہ مکرمہ میں بنیادی احکامات سکھائے جا رہے تھے اس لیے اس کا حکم ان صورتوں میں زیادہ ہے جبکہ مدنی سورتوں میں اخلاقیات، اداب جیسے موضوعات زیادہ ہیں۔ اس لفظ کا استعمال سب سے زیادہ سورۃ البقرہ اور اس کے بعد سورۃ النساء میں ہوا ہے۔



Graphical representation of IF-IDF score of the word “as-salaat” in maki and madni surrah

درستگی کا خلاصہ

- ٹی ایف-ائی ڈی ایف میں الفاظ کی تعداد کی درستگی 75٪ رہی ہے
- ی ایف-ائی ڈی ایف موضوعاتی درجہ بندی 20٪ کیونکہ وہ مانوی سیاق وسباق کو پروسیس کرنے کی توانائی نہیں رکھتا
- ٹی ایف-ائی ڈی ایف مشترکہ حروف کی نشاندہی 65٪
- ارابیرٹ کی موضوعاتی درجہ بندی 80٪

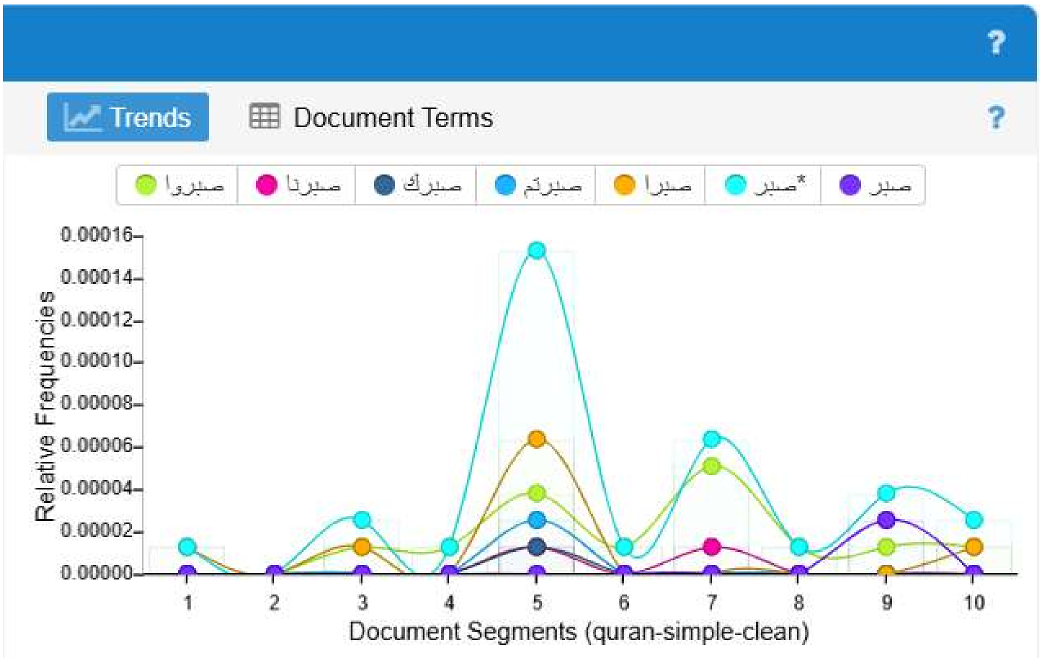
کیس اسٹڈی مکی اور مدنی سورتوں میں "صبر":

صبر صرف صفت نہیں بلکہ مومن کا ہتھیار ہے۔دین اسلام زندگی کے مختلف معاملات میں صبر کرنے کا درس دیتا ہے جیسے نیک کاموں پر جمے رہنا،برائی سے خود کو روکنا ، پریشانی اور مصیبت میں صبر کرنا۔ ارشاد باری تعالیٰ ہے:

"وَاسْتَعِينُوا بِالصَّبْرِ وَالصَّلَاةِ ۚ وَإِنَّهَا لَكَبِيرَةٌ إِلَّا عَلَى الْخَاشِعِينَ"[7]

"صبر اور نماز کے ساتھ مدد طلب کرو،یہ چیز شاق ہے،مگر ڈر رکھنے والوں پر۔"

قرآن پاک میں صبر کی مشتقات 44 سورتوں میں 103 بار آیا ہے۔ ان 44 سورہ میں سے 34 مکی اور 10 مدنی ہیں۔قرآن پاک میں یہ مختلف طرح سے استعمال ہوا ہے جیسے صبر، صبروا، صبرتم،، صبرک، صبرا۔



Graphical representation of IF-IDF score of the word "sabr" in maki and madni surrah

ٹی ایف-ائی ڈی ایف کے مطابق اس کا سکور مکی صورتوں میں زیادہ جبکہ مدنی سورتوں میں کم ہے۔بنیادی تفاسیر کے مطالعہ سے معلوم ہوتا ہے کہ آغاز اسلام میں مسلمانوں کو بہت سے مسائل،مشکلات اور ظلم وستم کا سامنا کرنا پڑا۔13 سال تک کفار مکہ نے مسلمانوں پر ہر قسم کا ظلم وستم کیا جس میں جسمانی تکالیف،معاشی معاشرتی بائیکاٹ ذہنی اذیت اور وطن چھوڑنے کا غم شامل ہے۔ لیکن ان مشکل حالات میں اللہ اور اس کے رسول صلی اللہ علیہ والہ وسلم نے تمام آزمائشوں پر مسلمانوں کو صبر اور ضبط نفس کا حکم دیا۔

ٹی ایف-ائی ڈی ایف کی وسعتِ گرفت اور نظریاتی حدود:

ٹی ایف ائی ڈی ایف صورتوں میں لفظی اصطلاحات کو سمجھ لیتا ہے لیکن وہ اصلاحات جنہیں علماء موضوعاتی کلیدی الفاظ کہتے ہیں یہ اس کی پہچان نہیں کر سکتا۔

معرفتی حد:

یہ تحقیق واضح طور پر اعلان کرتی ہے کہ کمپیوٹر یشنل الگور تھم علمی اساس نہیں بلکہ ایک تجزیاتی آلہ ہے۔ فہم القرآن کے کئی پہلو ہمیشہ الگور تھم کے دسترس سے باہر رہیں گے جیسے:

اسباب نزول: الگور تھم وحی کا سیاق وسباق نہیں سمجھ سکتے۔

ناسخ و منسوخ: ناسخ و منسوخ آیات میں فرق نہیں کر سکتے۔

اعجاز القرآن: قرآن پاک کی معجزانہ زبان کا مکمل فہم حاصل نہیں کر سکتے۔

احکام اور متشابہات: اس کا علم صرف علماء جانتے ہیں شماریاتی تجزیہ نہیں کر سکتے۔

حضور اکرم صلی اللہ علیہ وسلم کا ارشاد ہے:

" إِنَّ اللَّهَ لاَ يَقْبِضُ العِلْمَ انْتِزَاعًا يَنْتَزِعُهُ مِنَ العِبَادِ،وَلَكِنْ يَقْبِضُ العِلْمَ بِقَبْضِ العُلَمَاءِ "[8]

"اللہ علم کو اس طرح نہیں اٹھالے گا کہ اس کو بندوں سے چھین لے۔ بلکہ وہ علماء کو موت دے کر علم کو اٹھائے گا۔"

خلاصہ بحث(Conclusion)

اس مطالعے نے قرآنی متنی تجزیے میں TF-IDF اور ARABERT کے اطلاق کا جائزہ لیا گیا اور کلاسیکی تفسیر کے معیار پر ان کی درستگی پرکھی گئی۔ درج بالا نتائج محدود امید کی تصویر پیش کرتے ہیں: TF-IDF لفظی تعدد کاموں میں ۷۵-۸۰٪ اور موضوعاتی درجہ بندی میں 26-25٪ مطابقت رکھتا ہے، جبکہ ARABERT کی معنوی صلاحیت 80-88٪ درستگی کا تخمینہ لگاتی ہے اور مخلوط فریم ورک 84-89٪ تک پہنچتا ہے۔ یہ کمپیوٹیشنل تحقیق کے لیے بامعنی اعداد ہیں لیکن یہ ہمیں یاد دلاتے ہیں کہ 12-38٪ جو الگور تھم نہیں پکڑ سکتے، وہ صدیوں کا منتقل شدہ قرآنی علم ہے۔ اسبابِ نزول کی وضاحت کرنے والی احادیث، متشابہ آیات پر علمی اتفاقِ رائے (اجماع)، اور قرآنی معنی کے وہ روحانی پہلو جنہوں نے چودہ صدیوں سے مسلمانوں کی رہنمائی کی ہے۔ کوئی بھی الگور تھم، خواہ وہ کتنا ہی جدید کیوں نہ ہو، اس زندہ روایت کا متبادل نہیں بن سکتا۔

کمپیوٹیشنل ٹولز اس کام کو آسان، تیز اور معاون بنانے میں اہم کردار ادا کرتے ہیں۔ یہ پورے قرآن پاک کو سیکنڈوں میں پروسیس کر سکتا ہے شماریاتی نمونے پیش کر سکتا ہے اس کا درست استعمال قرآنی تحقیق کو مزید تقویت بخش سکتا ہے۔ اللہ کا یہ وعدہ ہے کہ اس نے قرآن کریم کو ذکر اور تدبر کے لیے آسان بنایا ہے لہذا اس کو سمجھنے کے لیے ہر جائز طریقہ ذکر و تدبر میں شامل ہے بشمول وہ کمپیوٹیشنل اجزاء جو جدید علم فراہم کرتے ہیں۔

نتائج وسفارشات:

- اس جدید دور میں اسلامیات کے طلبہ اور طالبات کے لیے ضروری ہے کہ سائنس اور ٹیکنالوجی کا بغور مطالعہ کریں اور اس سے فائدہ اٹھائیں۔
- ہم کمپیوٹیشنل آلات کا استعمال کرتے ہوئے اپنے کام کو انتہائی کم وقت میں مکمل کر سکتے ہیں۔
- بنیادی اساس سے کٹ کر حقیقی علم کا حصول ناممکن ہے۔ جدید آلات بنیادی کتب کا نعم البدل نہیں ہو سکتے۔

- اس بات کی ضرورت ہے کہ اسلامیات کے طلبہ اور طالبات کو کثیر الشعبہ مضامین پڑھائے جائیں کیونکہ ان سب کی بنیاد مسلمان سائنسدان اور مفکرین نے رکھی ہے۔

حوالہ جات

[1] القرآن(38:29)۔

[2] محمد بن اسماعیل بخاری، الجامع الصحیح (بیروت: دارالمغرب، 1992)، حدیث نمبر 5027۔

[3] القرآن(2:31)۔

[4] القرآن(2:23)-

[5] محمد بن عیسی ترمذی، الجامع الترمذی (بیروت: دار الغرب، 1998)، حدیث نمبر: 2950۔

[6] ترمذی، الجامع الترمذی، حدیث نمبر 2621۔

[7] القرآن(2:45)۔

[8] البخاری، الصحیح البخاری، حدیث نمبر 100۔