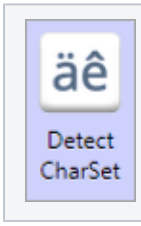
# Detect CharSet




**Detect CharSet**

**Author: Jerry Chae**

This plugin takes and text file and tells you what Character Set (Character Code /Encoding) the file uses.

Need help?

**Technical** contact to tech@argos-labs.com

**May you search all operations,**

- **Actions**
- **Verifications**
- **System Calls**
- **Interactives**

---

# Input

1. Any text files (.txt or .csv are the more common ones)

## Output

1. A CSV file with 3 headers

   Headers are: Language, Encoding, Confidence.

| | A | B | C | D |
|---|---|---|---|---|
| 1 | language | encoding | confidence | |
| 2 | | utf-8 | 0.99 | |
| 3 | | | | |

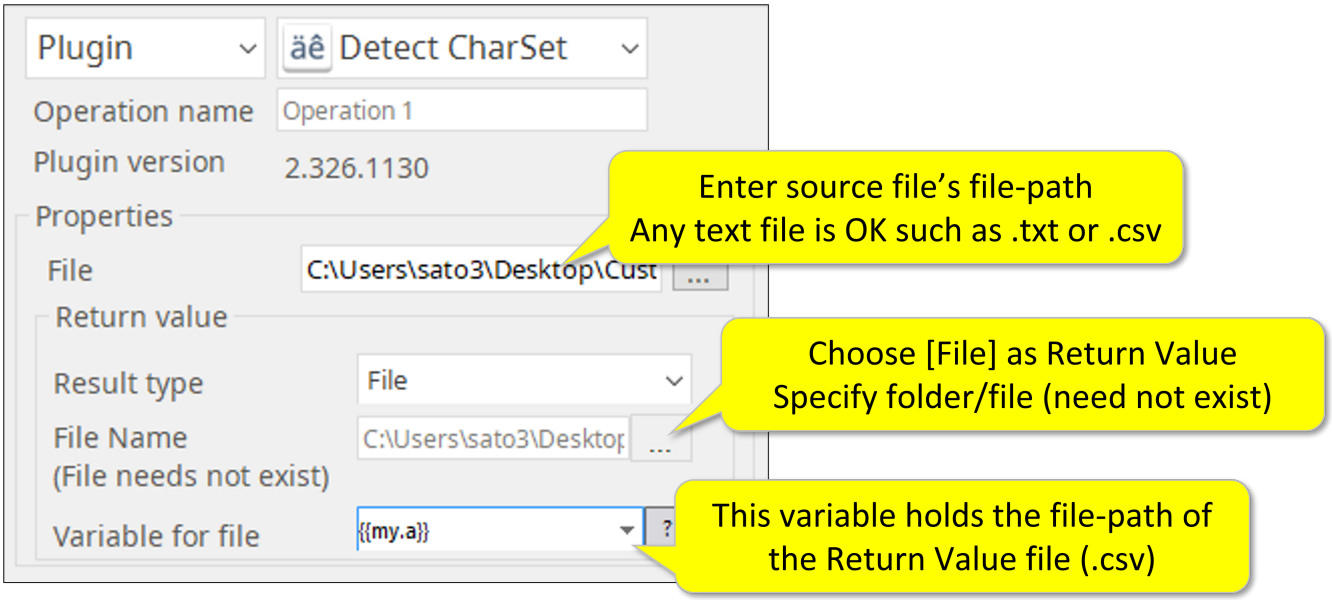
⚠️ Note: it reply blank when the language is for all languages.

## Character Code Table

| Codec | Languages | Codec | Languages | Codec | Languages | Codec | Languages |
|---|---|---|---|---|---|---|---|
| **ascii** | English | **cp869** | Greek | **gbk** | Unified Chinese | **johab** | Korean |
| **big5** | Traditional Chinese | **cp874** | Thai | **gb18030** | Unified Chinese | **koi8_r** | Russian |
| **big5hkscs** | Traditional Chinese | **cp875** | Greek | **hz** | Simplified Chinese | **koi8_t** | Tajik |
| **cp037** | English | **cp932** | Japanese | **iso2022_jp** | Japanese | **koi8_u** | Ukrainian |
| **cp273** | German | **cp949** | Korean | **iso2022_jp_1** | Japanese | **kz1048** | Kazakh |
| **cp424** | Hebrew | **cp950** | Traditional Chinese | **iso2022_jp_2** | Japanese, Korean, Simplified Chinese, Western Europe, Greek | **mac_cyrillic** | Bulgarian, Byelorussian, Macedonian, Russian, Serbian |
| **cp437** | English | **cp1006** | Urdu | **iso2022_jp_2004** | Japanese | **mac_greek** | Greek |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **cp500** | Western Europe | **cp1026** | Turkish | **iso2022_jp_3** | Japanese | **mac_iceland** | Icelandic |
| **cp720** | Arabic | **cp1125** | Ukrainian | **iso2022_jp_ext** | Japanese | **mac_latin2** | Central and Eastern Europe |
| **cp737** | Greek | **cp1140** | Western Europe | **iso2022_kr** | Korean | **mac_roman** | Western Europe |
| **cp775** | Baltic languages | **cp1250** | Central and Eastern Europe | **latin_1** | Western Europe | **mac_turkish** | Turkish |
| **cp850** | Western Europe | **cp1251** | Bulgarian, Byelorussian, Macedonian, Russian, Serbian | **iso8859_2** | Central and Eastern Europe | **ptcp154** | Kazakh |
| **cp852** | Central and Eastern Europe | **cp1252** | Western Europe | **iso8859_3** | Esperanto, Maltese | **shift_jis** | Japanese |
| **cp855** | Bulgarian, Byelorussian, Macedonian, Russian, Serbian | **cp1253** | Greek | **iso8859_4** | Baltic languages | **shift_jis_2004** | Japanese |
| **cp856** | Hebrew | **cp1254** | Turkish | **iso8859_5** | Bulgarian, Byelorussian, Macedonian, Russian, Serbian | **shift_jisx0213** | Japanese |
| **cp857** | Turkish | **cp1255** | Hebrew | **iso8859_6** | Arabic | **utf_32** | all languages |
| **cp858** | Western Europe | **cp1256** | Arabic | **iso8859_7** | Greek | **utf_32_be** | all languages |
| **cp860** | Portuguese | **cp1257** | Baltic languages | **iso8859_8** | Hebrew | **utf_32_le** | all languages |
| **cp861** | Icelandic | **cp1258** | Vietnamese | **iso8859_9** | Turkish | **utf_16** | all languages |
| **cp862** | Hebrew | **cp65001** | Windows only Windows UTF-8 (CP_UTF8) | **iso8859_10** | Nordic languages | **utf_16_be** | all languages |
| **cp863** | Canadian | **euc_jp** | Japanese | **iso8859_11** | Thai languages | **utf_16_le** | all languages |
| **cp864** | Arabic | **euc_jis_2004** | Japanese | **iso8859_13** | Baltic languages | **utf_7** | all languages |
| **cp865** | Danish, Norwegian | **euc_jisx0213** | Japanese | **iso8859_14** | Celtic languages | **utf_8** | all languages |
| **cp866** | Russian | **euc_kr** | Korean | **iso8859_15** | Western Europe | **utf_8_sig** | all languages |
| | | **gb2312** | Simplified Chinese | **iso8859_16** | South-Eastern Europe | | |

## How to set parameters



**All Plugins**

- ABBYY Download
- ABBYY Status
- ABBYY Upload
- AD LDAP
- Adv Send Email
- API Requests
- ARGOS API
- Arithmetic Op
- ASCII Converter
- Attach Image
- AWS S3
- AWS Textra Rekog
- Base64
- Basic Numerical Operations
- Basic String Manipulation
- Bot Collabo
- Box
- Box II
- Chatwork GetMessage
- Chatwork Notification
- Citizen Log
- Clipboard
- Codat API
- Convert CharSet
- Convert Image
- Convert Image II
- Create Newfile
- CSV2XLSX
- Data Plot I
- DeepL Free
- Detect CharSet
- Dialog Calendar
- Dialog Error
- Dialog File Selection
- Dialog Forms
- Dialog Info
- Dialog Password
- Dialog Question
- Dialog Text Entry
- Dialog Text Info
- Dialog Warning
- DirectCloud API
- Doc2TXT
- DocDigitizer Get Doc
- DocDigitizer Tracking
- DocDigitizer Upload
- Docker Remote Service
- Drag and Drop
- Dropbox
- Dynamic Python
- Email IMAP ReadMon
- Email Read Mon
- Env Check
- Env Var
- Excel2Image
- Excel Advanced
- Excel Advance IV
- Excel AdvII
- Excel AdvIII
- Excel Copy Paste
- Excel Formula
- Excel Large Files
- Excel Macro
- Excel Newfile
- Excel Simple Read
- Excel Simple Write
- Excel Style
- Excel Update
- Fairy Devices mimi AI
- File Conv
- File Downloader
- File Folder Exists
- File Folder Op
- File Status

- Fixed Form Processing
- Floating Form Processing
- Folder Monitor
- Folder Status
- Folder Structure
- FTP Server
- Git HTML Extract
- Google Calendar
- Google Cloud Vision API
- Google Drive
- Google Search API
- Google Sheets
- Google Token
- Google Translate
- Google TTS
- GraphQL API
- Html Extract
- HTML Table
- IBM Speech to Text
- IBM Visual Recognition
- Java UI Automation
- JP Holiday
- JSON Select
- JSON to from CSV
- Lazarus Forms
- Lazarus Invoices
- Lazarus RikAI
- Lazarus Riky
- LINE ID Card OCR
- LINE Notify
- LINE Receipt OCR
- Microsoft Teams
- MongoDB
- MQTT Publisher
- MS Azure Text Analytics
- MS Word Extract
- NAVER OCR
- Newuser-SFDC
- OCI
- OCR PreProcess
- OpenAI API
- Oracle SQL
- Outlook
- Outlook Email
- PANDAS I
- pandas II
- pandas III
- PANDAS profiling
- Parsehub
- Password Generate
- Path Manipulation
- PDF2Doc
- PDF2Table
- PDF2TXT
- PDF Miner
- PDF SplitMerge
- PostgreSQL
- PowerShell
- PPTX Template
- Print 2 Image
- Python Selenium
- QR Generate
- QR Read
- RakurakuHanbai API
- Regression
- Rename File
- REST API
- Rossum
- Running GAS
- Scrapy Basic
- Screen Capture
- Screen Recording START
- Screen Recording STOP
- Screen Snipping
- Seaborn Plot
- SharePoint
- Simple Counter
- Simple SFDC
- Slack

- Sort CSV
- Speed Test
- SQL
- SQLite
- SSH Command
- SSH Copy
- String Manipulation
- String Similarity
- Svc Check
- Sys Info
- Telegram
- Tesseract
- Text2PDF
- Text2Word
- Text Read
- Text Write
- Time Diff
- Time Stamp
- Web Extract
- Windows Op
- Windows Screen Lock
- Win UI Control
- Win UI Text
- Word2PDF
- Word2TXT
- Word Editor
- Work Calendar
- XML Extract
- XML Manipulation
- Xtracta Get Doc
- Xtracta Tracking
- Xtracta Upload
- YouTube Operation
- ZipUnzip